UNIVERSAL PERSONAL COMMUNICATIONS

Wideband CDMA for Third Generation Mobile Communications

Tero Ojanperä
Ramjee Prasad
Editors

Artech House Publishers BOSTON • LONDON

transmission power. The decrease in the transmission power of a mobile station as a function of multiuser detection efficiency is calculated analytically.

7.2 DIVERSITY

Diversity is essential in achieving good performance in a fading channel. With wideband CDMA, the means for achieving diversity are multipath, antenna, polarization, macro, and time diversity. The use of frequency hopping typical of TDMA systems, like GSM, is not employed with wideband CDMA to gain frequency diversity.

7.2.1 Gain From Multipath Diversity

The uncoded bit error rate as a function of SNR and the number of diversity branches is given as [1]

$$P_b = \left[\frac{1}{2}(1-\mu)\right]^L \sum_{k=0}^{L-1} \binom{L-1+k}{k} \left[\frac{1}{2}(1+\mu)\right]^k \tag{7.1}$$

where L is the number of diversity branches and

$$\mu = \sqrt{\frac{\bar{\gamma}_c}{1+\bar{\gamma}_c}} \tag{7.2}$$

where $\bar{\gamma}_c$ is the average SNR per diversity branch. It is assumed that all diversity branches have equal SNR on average.

The importance of diversity can be seen in Figure 7.1, where the average uncoded BER is plotted as a function of the total SNR in a Rayleigh fading channel for PSK modulation with Lth order diversity. Ideal coherent maximal ratio combining of uncorrelated diversity branches is assumed. Maximal ratio combining requires estimation of the amplitudes and phases of the diversity components. When the number of multipath components increases, the SNR per each multipath gets lower, and it becomes more difficult to perform coherent combining in a RAKE receiver. Also, when the diversity branches are correlated, the diversity gain gets lower.

7.2.2 Multipath Diversity in Different Radio Environments

One of the key factors that differentiate the third generation wideband CDMA from the second generation narrowband CDMA is the wider bandwidth. In addition to the ability to provide wideband services, the increased bandwidth makes it possible to resolve more multipath components in a mobile radio channel. If the transmission bandwidth is wider than the coherence bandwidth of the channel, the receiver is able to separate

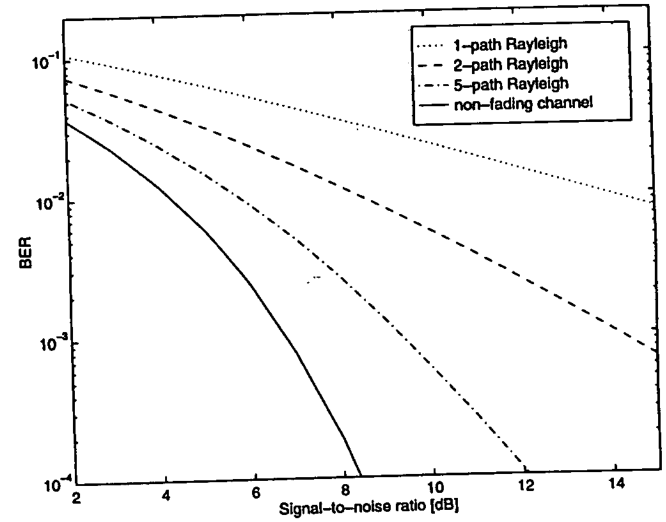


Figure 7.1 Uncoded BER with diversity assuming ideal coherent combining of diversity branches.

multipath components. This brings more diversity and higher capacity, along with power control.

We already discussed the effect of bandwidth on the diversity order in Chapter 4, where ATDMA and CODIT channel models were used. In this section, we use the ITU channel models presented in [2-3]. ITU channel models are based on the JTC (Joint Technical Committee for the U.S. PCS standardization) and ATDMA models [2]. Table 7.1 presents the coherence bandwidths for different environments. Also, the theoretical number of separable paths with chip rates of 1.2288, 4.096, and 8.192 Mcps is calculated.

The coherence bandwidth is calculated as in [4]:

$$(\Delta f)_c = \frac{1}{2\pi S} \tag{7.3}$$

where S is the rms delay spread defined as in [5]:

$$S = \sqrt{\frac{\int_0^\infty (\tau - D)^2 P(\tau)\,d\tau}{\int_0^\infty P(\tau)\,d\tau}} \tag{7.4}$$

and $P(\tau)$ is the power delay profile, and D is the average delay defined as

$$D = \frac{\int_0^\infty \tau P(\tau)d\tau}{\int_0^\infty P(\tau)d\tau} \tag{7.5}$$

The maximum number of rays (paths) in the ITU channel models is 6.

Table 7.1
Coherence Bandwidths of ITU Channel Models

			Maximum number of separable paths		
ITU Channel model	Number of paths in channel model	Coherence bandwidth	1.25 Mcps	4.0 Mcps	8.0 Mcps
Indoor A	6	4.3 MHz	1	1	2
Indoor B	6	1.6 MHz	1	3	6
Outdoor to indoor A	4	3.5 MHz	1	2	3
Outdoor to indoor B	6	250 kHz	5	17	33
Vehicular A	6	430 kHz	3	10	20
Vehicular B	6	53 kHz	24	78	155

Source: [6].

With a chip rate of 1.2288 Mcps, only very limited multipath diversity is available in indoor and in outdoor-to-indoor environments, for ITU channel models.

In Table 7.2, the number of useful multipath components is presented with three different environments measured in Australia [7]. The measured results indicate clearly the superiority of the wideband CDMA compared to narrowband CDMA in terms of achieving diversity gain with the RAKE receiver. With the 1.25-MHz bandwidth only one or two multipath components are available, where a single finger collects most of the energy. Very limited multipath diversity can be achieved with 1.25-MHz. With the 4-MHz bandwidth the achieved number of multipath components is in all measurements either two or four, which means a clearly achievable diversity gain.

If the autocorrelation properties of the spreading codes are not ideal, multipath components interfere with each other degrading the multipath diversity gain. This effect will be noticeable with very low spreading factors (i.e., with high bit rates).

Table 7.2
Number of Useful Multipath Components

	1.25 MHz	4.0 MHz	8.0 MHz
Melbourne	1	2	2
Adelaide	1	2	3
Sydney	2	4	6

Source:[7].

7.2.3 Antenna Diversity

If antenna diversity branches are close to each other, they are correlated and diversity gain is smaller. Therefore, antenna diversity is best suited for base stations and for large sized mobile stations.

Receiver antenna diversity can be used to average out receiver noise in addition to providing diversity against fading and interference. At base stations, receiver antenna diversity can be used to increase cell coverage by increasing the noise limited uplink range.

Transmission antenna diversity can be used to generate multipath diversity in such environments where only limited multipath diversity is available (i.e., in indoor and macrocell). Transmission diversity can be obtained in several ways (See Section 5.11.2.1). The same signal can be transmitted from more than one antenna with a delay that the RAKE receiver can separate. Alternatively, the data can be divided between transmission antennas without any delays, thus maintaining the orthogonality of the diversity signals. Transmission antenna diversity is best suited for the downlink transmission in the base station together with the uplink diversity reception.

7.2.4 Polarization Diversity

Different polarization directions may experience different fading. Especially in indoor environments, polarization directions have been shown to be nearly uncorrelated, thus providing diversity [8]. The advantage of polarization diversity over antenna diversity is that polarization diversity does not require separation between antennas, and thus, it could be applied to small sized equipment as well.

7.2.5 Macro Diversity

With CDMA systems, the use of macro diversity (i.e., soft handover) is essential for achieving reasonable system performance because of frequency reuse 1 and fast power control. If the mobile station is not connected to the base station to which the attenuation is the lowest, unnecessary interference is generated in adjacent cells. In the uplink, the macro diversity effects are only positive, since the more base stations try to detect the signals, the higher the probability is for at least one to succeed. In the uplink direction, the detection process itself does not utilize the information from the other base stations receiving the same signal, but the diversity is selection diversity, where the best frame is selected in the network based on the frame error indication from a CRC check. The selection can be done in the base station controller (BSC) or in some other network element, depending on the implementation.

In the downlink, macro diversity is different as the transmission now originates from several sources and diversity reception is handled by one receiver unit in the mobile station. All extra transmissions contribute to the interference. Capacity improvement is based on a similar principle as with a RAKE receiver in a multipath channel, where the received power level fluctuations tend to decrease as the number of separable paths increases. With downlink macro diversity, the RAKE receiver

capability to gain from the extra diversity depends also on the number of available RAKE fingers. If a RAKE receiver is not able to collect enough energy from the transmissions from two or, in some cases, three base stations due to a limited number of RAKE fingers, the extra transmissions to the mobile can have a negative effect on the total system capacity due to increased interference. This is most likely in the macro cellular environments because the typical number of RAKE fingers considered adequate for capturing the channel energy in most cases is four. If all connections offered that amount of diversity, then the receiver has only one or two branches to allocate for each connection.

Table 7.3 illustrates the gain of soft handover over hard handover in the downlink. The results are generated with the system simulator as presented later in Section 7.3.2. The CODIT macro- and microcell environments, described in Chapter 4, were used in simulations [9].

From the macrocell results it can be seen that system downlink capacity decreases by 10% if macro diversity is not utilized. Capacity loss is not remarkable even if soft handover gain and macro diversity gain is lost. A CODIT macrocell channel gives sufficient diversity even if the mobile station was in a hard handover state. On the other hand, the limited number of RAKE fingers cannot fully exploit the attained diversity from soft handover in the macrocellular environment. On the contrary, a remarkable part of the energy generated by soft handover base stations is lost, which contributes to the interference. The situation is different for microcell, since a CODIT microcell channel provides only little diversity. Thus, macro diversity is essential for high capacity in environments with a low diversity order.

Table 7.3
Downlink Capacity with Hard and Soft Handover

[Kbps/cell/MHz]	Hard handover	Soft handover
Macro	155	169
Micro	139	222

In addition to providing gain against multipath fading, macro diversity also gives gain against shadowing. The shadow fade margin for soft handover is analyzed to be 2 to 3 dB smaller than the fade margin for hard handover [10,11]. Macro diversity can thus be used to increase the cellular range.

Soft handovers should be used with circuit switched services since macro diversity is important for low delay services to guarantee high quality with short delays. For packet data there are no strict delay requirements, and therefore, soft handover is not that important for nonreal-time packet data services that can take advantage of time diversity through retransmissions.

7.2.6 Time Diversity

Time diversity is achieved by coding, interleaving, and by retransmissions. Channel coding is applied to achieve lower power levels and required signal quality in terms of BER/FER. Interleaving and channel coding processes are used to correct the errors due to channel fades and interference peaks. The longer the interleaving, the better an error burst is spread over the interleaving period. Consequently, the individual errors can be much better handled with error correction coding than when errors appeared to the decoder in bursts. In practice, it is the delay requirements that limit the allowed interleaving depth. For nonreal-time services the requirements are clearly looser, and retransmissions can be employed to ensure near error-free transmission of data. The total delay requirement for real-time services is typically on the order of 20 to 100 ms.

As the interleaving period gets shorter, the requirements for fast power control get tighter. Generally with CDMA signal transmission there is a peak in the BER curve when plotting the BER as a function of mobile speed; until certain velocities, the power control eliminates the fades and, at some speed, the power control is not able to follow the fades anymore [12–13]. But as the fades occur more often and get shorter in the time domain, the interleaving is able to spread the information over several fades and thus uncorrelate the signal errors to allow them to be corrected with forward error correction such as convolutional coding. However, when antenna diversity is applied, the performance is almost flat over different mobile speeds [12].

7.3 W-CDMA SIMULATORS

In this section the principles of the evaluation of the WCDMA radio network performance are analyzed. There are two approaches to simulate the overall performance of a DS-CDMA system. One is a combined approach where the link level and cellular network level simulations are combined into one simulator. Another approach is to separate the link and system level simulations to reduce the complexity of the simulators. For the performance evaluation of cellular networks a single simulator approach would be preferred. The complexity of such simulator — including everything from transmitted waveforms to a cellular network with tens of base stations — is far too high. Therefore, separate link and system level simulators are needed.

In the link level, one communication link between a mobile station and a base station is modeled. The time resolution is typically 1 to 4 samples per chip. For 4.0 Mcps with one sample per chip, the time resolution is 0.25 μs. In the system level there are tens of base stations and all the mobiles connected to those base stations. The time resolution is coarser than in the link level, typically one power control period (0.5 ms) or longer.

The interface between the link level simulator and the system level simulator is presented in Figure 7.2 and in [14]. The link level simulator provides the system level simulator with the required input parameters. Those parameters are E_b/N_0 both for uplink and downlink, MUD efficiency for uplink, orthogonality factor for downlink, and multipath channel model including antenna diversity. Here, E_b is energy per received bit and N_0 is the interference power per bit. The orthogonality factor could be also calculated theoretically based on the multipath profile.

MUD efficiency states the percentage of intracell interference that can be cancelled with base station multiuser detection. In system level simulations it is assumed that MUD efficiency is the same for different system loads.

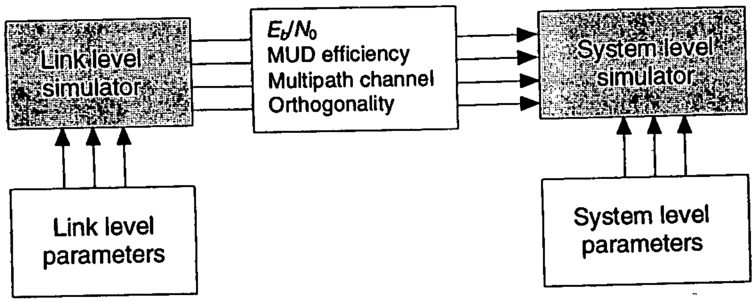


Figure 7.2 The interface between the link level and the system level simulator.

7.3.1 Link Level Simulation

The simulation program used in the link level performance evaluations was the Communications Simulation and System Analysis Program (COSSAP) by Synopsys Inc. The COSSAP simulation program is a stream driven simulation program, with support for asynchronous operation since no timing control between simulation elements is required. All the receiver blocks were generated with the C language for efficiency and flexibility in the configuration.

Link level simulations included a multipath channel model that was modeled as a tapped delay line with one sample per chip. The CODIT macro-, micro-, and picocell channel models described in Chapter 4 were used. Amplitudes and phases of the multipath components were estimated in the receiver, but the multipath delays were known a priori. Fast power control in the link level simulations was based on the estimated received SIR values.

7.3.2 System Level Simulation

7.3.2.1 *Simulator Principle*

The system level capacity is studied in an interference-limited case. One capacity simulation typically contains 10,000 local mean SIR values. One local mean SIR value is simulated over a 500-ms period for each user, consisting of random mobile placement, virtual handover, and fast power control. Figure 7.3 presents the block diagram of the system level simulator. The simulations give the outage probability for the system, where outage probability represents the number of users having a worse average SIR than required. An outage probability of 5% was selected as the target value for capacity.

The mobile stations in this simulator are not actually moving, such as in dynamic simulators with mobility. The effect of mobility is modeled by adding a fast

fading process to the signal in the system level. In this semistatic approach, signal strength varies due to fading as a function of the mobile speed given as a parameter. In the system level it is important to model the effect fast fading and fast power control as discussed in Section 7.4. The simulator can also be used for multioperator simulations.

The propagation model of the system simulator consists of attenuation, shadowing, and statistically generated fast fading. Both macro- and microcell models are adopted from [15].

7.3.2.2 *Path Loss Models*

The distance-dependent attenuation for macro- and microcell environments is used as presented in Section 4.5 for macro- and microcell environments. The simulation environment parameters are given in Table 7.4.

The path loss between a mobile station and a macrocell base station is modeled as

$$L = 29 + 36\log(R) + 31\log(f) \tag{7.6}$$

where f is the carrier frequency in megahertz and d the distance between the transmitter and the receiver in kilometers. Resulting path loss L is given in decibels.

Path loss between a mobile station and a microcell base station is calculated with a multislope model. The microcell base stations are located on every second street intersection. The slopes are non-line-of-sight slope L_{NLOS}, line-of-sight slope L_{LOS} for small distance and line-of-sight slope for long distances. If the connection between transmitter and receiver is line of sight, attenuation is calculated as

$$L_{LOS} = \begin{cases} 82 + 20\log(\frac{d}{300}), \; if \quad d \le 300\text{m} \\ 82 + 40\log(\frac{d}{300}), \; if \quad d > 300\text{m} \end{cases} \tag{7.7}$$

At a distance of 300m a breakpoint marks the separation between two line-of-sight segments.

Turning a corner causes an additional loss in (7.7). Attenuation between a transmitter and a receiver that have non-line-of-sight connections constitutes a line-of-sight segment, a non-line-of-sight segment, and an additional corner attenuation.

$$L_{NLOS} = L_{LOS}(d_{corner}) + 17 + 0.05d_{corner} + (25 + 0.2d_{corner})\log\left(\frac{d}{d_{corner}}\right) \tag{7.8}$$

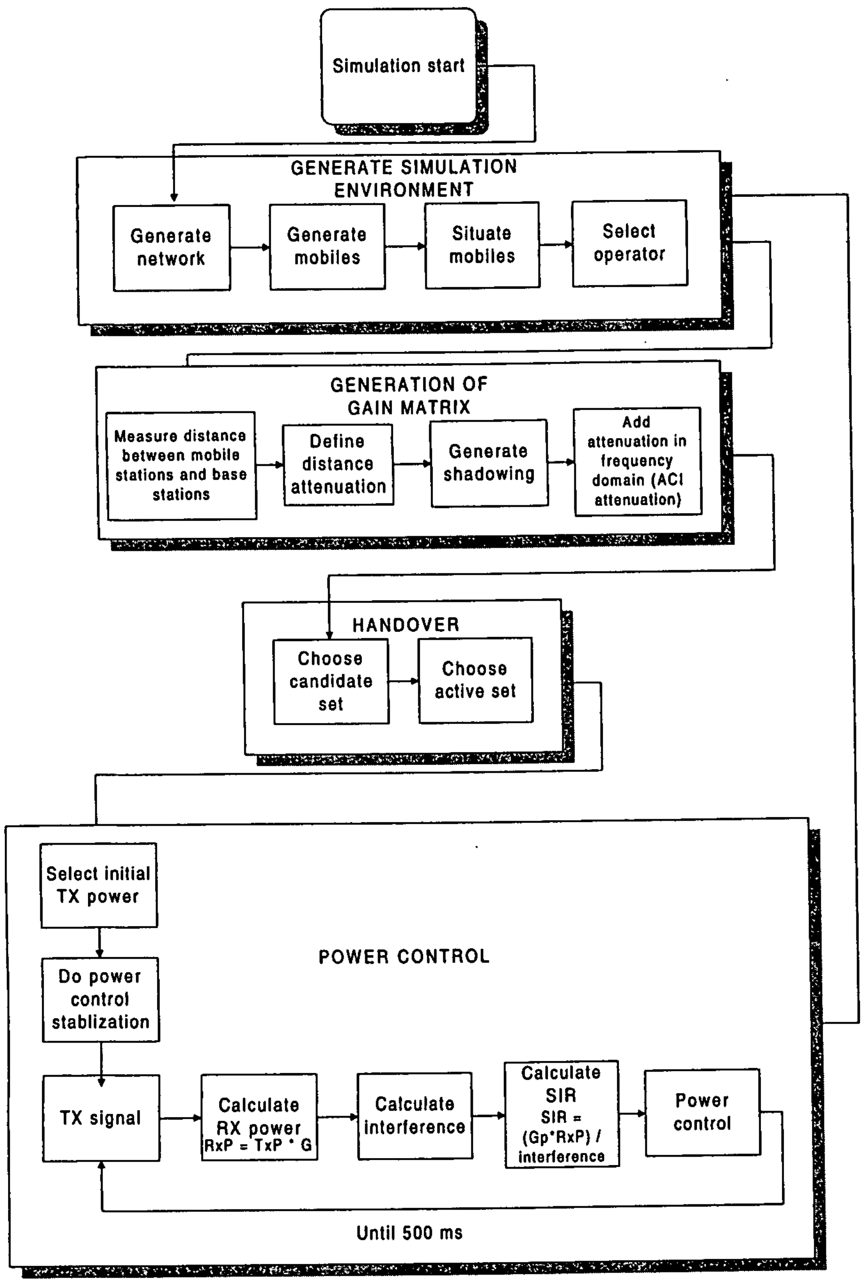


Figure 7.3 The block diagram of the system level simulator.

Line-of-sight attenuation is calculated between a corner and receiver d_{corner}, and non-line-of-sight connection between a corner and transmitter.

Table 7.4
Simulation Environment

	Micro	Macro
Shadowing	mean 0 dB, std dev 4dB	mean 0 dB, std dev 10dB
Number of base stations	128	19
Base station layout	Manhattan	hexagonal
Base station spacing	200m	800m
Building width	100m	-
Street width	30m	-
Mobile speed	3 km/h 36 km/h	5 km/h 50 km/h

Figures 7.4 and 7.5 depict, respectively, the macrocellular base station location and the base station deployment in microcellular system level simulations.

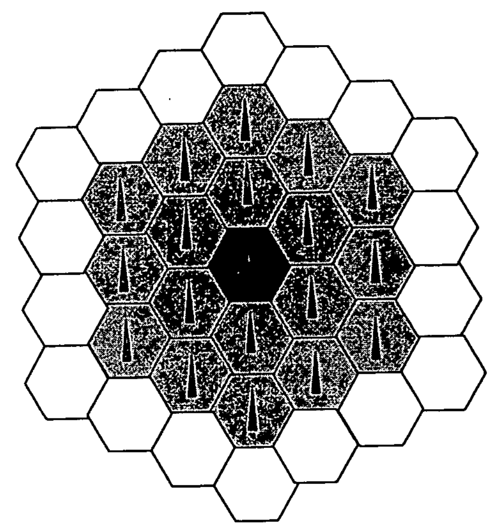


Figure 7.4 Macrocellular base station location.

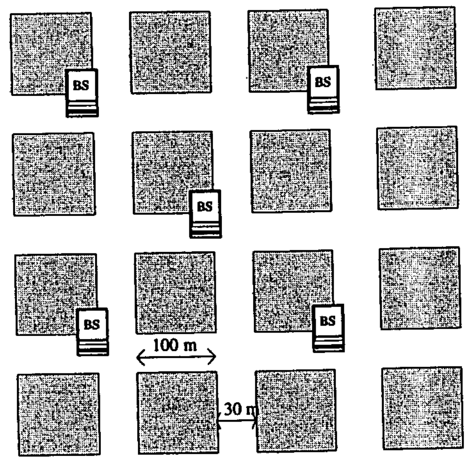


Figure 7.5 Base station deployment in a microcellular system level simulator.

7.3.2.3 *Fast Fading and Fast Power Control Modeling*

In order to correctly model CDMA effects such as fast power control, a fast fading functionality was included in the system simulator, with the principle presented in [4]. The same kind of multipath channel with Rayleigh fading multipath components was used in the system level as in the link level simulations.

The power control algorithm is implemented as in [16]. The signal-to-interference ratio (SIR)-based power control algorithm aims to keep SIR levels of users at an appropriate level by adjusting the transmission power up or down. The received SIR level is compared to the SIR threshold. If the received SIR was lower than the threshold, a "power up" command is sent. If the SIR level was higher than the threshold, a "power down" command is sent.

In the link level simulations, realistic fast power control has been modeled. Real SIR estimation has been used in the WCDMA receiver, and errors have been added to the power control commands in the feedback channel. Also, real power control frequency and step size have been used. The effect of nonideal power control can be seen in the obtained E_b/N_0 figures.

In the system level, ideal SIR estimation has been used in the fast power control since the effect of nonideal power control is taken into account in the link level simulations. In the system level, fast fading and fast power control are still needed because we can observe how power control effects the interference in the network. Therefore, real power control frequency and step sizes are used also in the system level.

7.3.2.4 *Handover Modeling*

In the handover modeling, a maximum of 3 BTSs are selected in the active set. The BTSs are selected randomly with a uniform distribution from a pool of BTSs that fits in the handover margin. Such a BTS selection procedure takes the effects of nonideal handovers into account.

In the uplink, macro diversity (i.e., soft handover) is taken into account by selecting the best source (the frame with highest average SIR) on a frame-by-frame basis. In the downlink, macro diversity can be modeled in two ways at the system level. The first approach is to simulate macro diversity in the link level so that separate E_b/N_0 levels are obtained for users in the soft handover state. There is a difference in E_b/N_0 levels because a different amount of diversity is available during soft handover and because part of the transmitted energy may not be captured by the limited number of RAKE fingers in the mobile station. Then, when we have separate E_b/N_0 values for the soft handover state in the system level, the mobile can be assumed to be able to receive all the paths directed to it in the system level.

Another way to model downlink macro diversity is to transmit the signal from several base stations to mobile stations so that the mobile station cannot receive all the paths directed to it. Paths not captured by RAKE processing contribute to the interference. Now the same E_b/N_0 threshold is used for all the mobile stations. The second method has been used in the system simulator due to straightforward modeling. The first method could be more accurate because it can also take into account the effect of increased multipath diversity on the E_b/N_0 performance.

7.3.2.5 *Micro Diversity Modeling*

Micro diversity (i.e., multipath diversity and antenna diversity in the BTS) is modeled by uncorrelated Rayleigh fading multipath components. The diversity combining is modeled as coherent maximal ratio combining. In this system, level micro diversity modeling, it is assumed that the number of RAKE fingers at the base station is sufficient and all modeled multipath components are received. The performance loss due to the limited number of RAKE fingers is modeled in the link level, and can be seen in the E_b/N_0 values.

Micro diversity modeling is needed in the system level because the amount of diversity affects the intercell interference, as shown in Section 7.4.2.

7.3.2.6 *Modeling Downlink Pilot and RAKE*

In the downlink, the pilot signal is essential for system operation, and it is generally transmitted with a higher power level than any traffic channel. Modeling of the common pilot is not included in the link level performance figures. For this reason, its effect on downlink interference must be included in the system level modeling. In the simulator, each base station transmits user-specific signals and a common pilot signal. Powers of all the transmitted signals are summed together at the base station transmitter. Coherent maximal ratio combining at the RAKE receiver is assumed, and it is modeled in the system level by taking the sum of SNR values of paths [4].

7.3.2.7 *Interference Cancellation*

Intracell interference cancellation is modeled in the system simulator by removing β percentage of intracell interference in the interference calculation. The factor β is called MUD efficiency, and it is obtained from the link level simulations shown in Section 7.5.1.

In the uplink, when measuring SIR for a single user, the cancellation process includes all the users connected to the same BTS. Interference cancellation is nonideal, with the efficiency derived from the link level simulations. Thus, the total interference that the BTS experiences consists of the interference from all the mobiles in the network not having a connection with the BTS and from the partly canceled interference from the mobiles having connection to the BTS.

7.3.2.8 *Sectorization and Adaptive Antennas*

The effects of sectorization are not included in this study. The capacity gain from sectorization is $\xi \cdot N$; where ξ is the sectorization efficiency, and N is the number of sectors. Coverage overlap between antenna beams causes interference spillover, which causes sectorization efficiency to be less than 1 and reduces capacity. For 3-dB overlap between antennas, sin(x)/x antenna pattern gives a sectorization efficiency of 0.88. When the beam overlap increases, the efficiency decreases. The efficiency also depends on whether softer handoff is used between overlapping sectors [17]. The use of adaptive antennas for improving capacity or coverage is not covered here, but evaluation results can be found in [18].

7.3.3 Simulation Parameters

The wideband CDMA system parameters used in the simulation are listed in Table 7.5. These system parameters are from the MUD-CDMA system concept [19], a wideband CDMA system that has been used as a basis for the FRAMES FMA2 wideband CDMA air interface [20]. Since the main system parameters are similar to the other wideband CDMA system presented in Chapter 6, the results are representative for third generation wideband CDMA systems.

7.4 FAST POWER CONTROL

As was discussed in Chapter 5, fast power control impacts the performance of a wideband CDMA system in three different ways. First, it compensates for the fast fading and, in a perfect case, turns the fading channel into a nonfading channel in the receiver, reducing the required E_b/N_0. The second effect is a consequence of the first: the compensation of the fading channel by power control leads to peaks in mobile station transmission power, which affect the intercell interference in the cellular network. And thirdly, fast power control equalizes the mobile station powers in the base station and thus prevents the near-far effect. The effect of fast power control on a CDMA system can be summarized as follows:

- Lowers the E_b/N_0 requirement in the receiver;
- Introduces interference peaks in the transmitter;
- Avoids near-far problems in the uplink.

Table 7.5
Simulation Parameters

CDMA System Parameters	
Carrier spacing	6 MHz
Chip rate	Basic chip rate 5.1 Mcps
Carrier frequency	2 GHz
Spreading codes	Uplink: Extended VL-Kasami Downlink: Augmented Walsh-Hadamard
Bit rates	12 Kbps 144 Kbps 2 Mbps
Interleaving	20 ms (12 Kbps) 40 ms (144 Kbps) 100 ms (2 Mbps)
PC dynamics	Uplink 80 dB Downlink 20 dB
Link Level Parameters	
Voice activity	100%
Number of stages in base station MUD	2 (1 interference canceling)
Antenna diversity (uplink)	2 antennas, with fading correlation 0.7 and 0.0
Path combining method	Maximal ratio combining
Target BER	10^{-3}
Coding	Uplink: 1/2 rate convolutional with K = 9 Downlink: 1/3 rate convolutional with K = 9
Fast power control	2 kHz, step size 1.0 dB, 5% errors in feedback signaling
RAKE fingers	4 or 9
System Level Parameters	
Outage requirement	95% satisfied users
Active set size in soft handover	3
Handover margin in nonideal handover	3 dB
PC step size	1 dB
Pilot strength	6 dB higher than maximum traffic channel power

In practice there are imperfections in the power control because of the transmission and processing delay, command errors, limited dynamic range, and step size; the system performance is degraded because of the increased residual variation in the signal-to-noise ratio.

7.4.1 Impact of Fast Power Control in E_b/N_0

The gain in E_b/N_0 depends on the channel characteristics, the bandwidth of the system, and the mobile speed. Since a wide bandwidth gives better diversity, power control is

not as critical for the wideband CDMA as for the narrowband CDMA. In a small delay spread channel, the improvement of performance from fast power control is largest. Furthermore, as discussed in Section 7.2.6, the improvement of E_b/N_0 is highest at slow mobile speeds. This is due to the ability of fast power control to follow the fast fading. As shown in Section 7.5.1 for a 2-Mbps service, fast power control is beneficial especially in micro- and picocell channels due to a limited multipath diversity. In a macrocell channel, the gain of fast power control in E_b/N_0 is not very large because of the high degree of multipath diversity. Fast power control is compared to slow power control in the downlink in [13,21].

7.4.2 Intercell Interference With Uplink Fast Power Control

When uplink fast power control is able to follow the fast fading, the intercell interference will increase due to peaks in the transmission (TX) power of the mobile MS in Figure 7.6. The received (RX) interference level in the neighboring cells experiences peaks. In the downlink, this effect is smaller due to limited power control dynamics. Downlink power control is not considered in this subsection. The less diversity available, the higher the average transmission power is and the more intercell interference is generated. This phenomenon of increased interference is imminent with such systems with fast power control that cannot exploit the multipath diversity of the channel and do not employ antenna diversity. On the other hand, soft handovers provide more diversity and reduce intercell interference. As explained below, if a system employs fast power control and soft handovers, they should be modeled together with a relevant channel model in the system simulations to give valid capacity results.

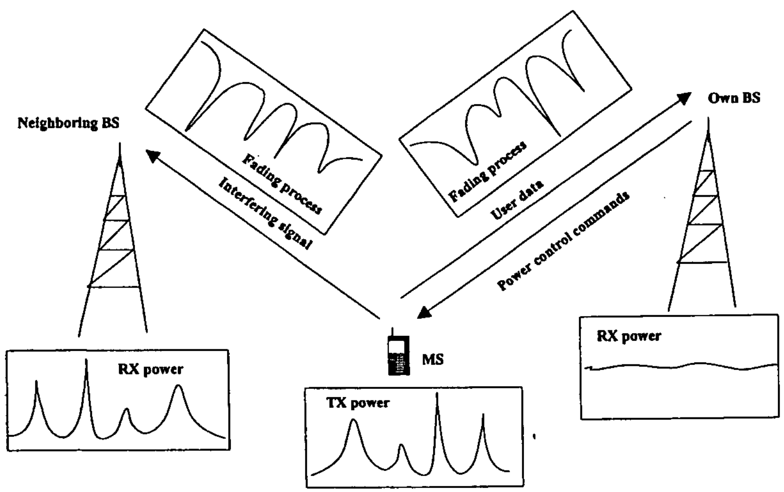


Figure 7.6 Interference to neighboring base stations in fading channel with fast power control without macro diversity.

In this section, we fix the SIR target to be the same for all channels in order to study only the impact of increased intercell interference. It should be noted that in practice this is not the case, but different channels require different SIR to achieve the required QoS.

In the following treatment, we first provide an analytical model and then the corresponding simulation results. Due to modeling difficulties, handovers are not considered in the analysis corresponding to the assumption of slow hard handovers, where a mobile is connected to the base station with the smallest long-term attenuation. This hard handover analysis leads to a worst case analysis in terms of transmission power and interference. In reality, soft handover increases diversity for mobile stations at the cell border, and thus, the interference is reduced. This is visible in the system simulations where both hard handover and soft handover are modeled.

7.4.2.1 *Average Uplink Transmission Power With Fast Power Control*

In this section an analysis on the average uplink transmission power with ideal fast power control is presented. The terminal transmission power level depends on the multipath diversity of the channel model. This result is then used in the next subsection to analyze system capacity with fast power control.

The average transmission powers with different degrees of diversity are calculated as follows. The probability density function (PDF) of the total channel power with L statistically independent Rayleigh fading multipath components, given in [1], is

$$p(\gamma)=\sum_{k=1}^{L}\frac{\pi_k}{\overline{\gamma_k}}e^{-\gamma/\overline{\gamma_k}}\qquad \gamma\geq 0 \tag{7.9}$$

where $\overline{\gamma_k}$ is the average power of the kth multipath component and

$$\pi_k=\prod_{\substack{l=1\\l\neq k}}^{L}\frac{\overline{\gamma_k}}{\overline{\gamma_k}-\overline{\gamma_l}} \tag{7.10}$$

If the multipath components are equally strong, the PDF is written as

$$p(\gamma)=\frac{L^L\gamma^{L-1}}{(L-1)!}e^{-L\gamma_0}\qquad \gamma\geq 0 \tag{7.11}$$

Now we want to find out the PDF of the transmission power ρ of a given mobile station. In this section, we assume that the number of interfering users is so high that the interference at the base station can be approximated by Gaussian noise. The more users there are in the system, the better this approximation holds. With this assumption, the interference level at the base station is constant. Therefore, SIR-based power control can be replaced by power control based on received signal power only. Since we

assume ideal power control, mobile transmission power $\rho = 1/\gamma$ and the PDF of ρ is [22]

$$p_\rho(\rho) = \frac{1}{\rho^2} p\left(\frac{1}{\rho}\right) \qquad \rho \geq 0 \tag{7.12}$$

The average transmission power can be obtained by

$$\overline{\rho} = \int_0^\infty \rho p_\rho(\rho) d\rho \tag{7.13}$$

In case of equally strong Rayleigh fading components and unlimited power control dynamics, the increase in transmission power can be shown to be

$$\frac{L}{L-1} \tag{7.14}$$

where L is the number of multipath components. In a 1-path Rayleigh fading channel ($L = 1$) the result would be infinity with unlimited power control dynamics. Typical total dynamics of uplink CDMA power control are 80 dB, which takes care not only of fast fading but also distance attenuation and slow fading. In these calculations only fast fading is taken into account, and therefore, the power control dynamics allowed for the fast power control is set to be 50 dB. The minimum and maximum transmission powers are chosen to be $\rho_{\min} = 0.2$ and $\rho_{\max} = 2.0\times10^4$. Now, the average transmission power can be calculated as

$$\overline{\rho} = \int_0^{\rho_{\min}} \rho_{\min} p_\rho(\rho) d\rho + \int_{\rho_{\min}}^{\rho_{\max}} \rho p_\rho(\rho) d\rho + \int_{\rho_{\max}}^{\infty} \rho_{\max} p_\rho(\rho) d\rho \tag{7.15}$$

The average transmission powers are calculated in the AWGN channel, in the 1-tap Rayleigh fading channel, in the 2-tap Rayleigh fading channel (2 equally strong taps on average), in the 4-tap Rayleigh fading channel (4 equally strong taps on average), and in the ATDMA macrocell channel with and without base station antenna diversity. The 2-tap and 4-tap Rayleigh fading channels, with all taps equally strong, can be considered either as multipath channels or as antenna diversity. The impulse response of the ATDMA macrocell channel model used is shown in Figure 7.7. The used bandwidth is chosen to be so high that all the multipath components in the ATDMA macrocell channel model can be resolved. The minimum bandwidth needed for separating all the paths in the ATDMA macrocell channel is approximately 3.5 MHz. Maximal ratio combining of multipath components is assumed in this analysis.

The calculated increase in average transmission power compared to nonfading channel ΔP_{tx} is shown in Table 7.6.

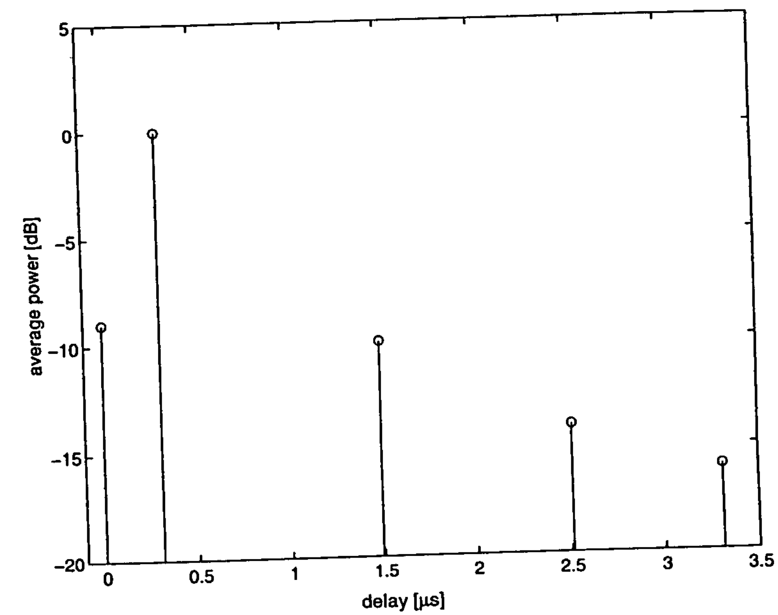


Figure 7.7 Impulse response of ATDMA macrocell channel.

Table 7.6
Average Increase of Transmission Power in a Fading Channel Compared to a Nonfading Channel With Hard Handovers

Channel	ΔP_{tx}(dB)
Nonfading channel	0
1-tap Rayleigh fading channel	10.1
2-tap Rayleigh fading channel	3.0
4-tap Rayleigh fading channel	1.2
ATDMA macro channel	2.5
ATDMA macro channel with antenna diversity	1.3

The average transmission power in a fading channel with fast power and with hard handovers is higher than in a nonfading channel. Thus, the average intercell interference from the mobile stations to the base stations in the surrounding cells is also higher.

7.4.2.2 *Analysis of Impact of Fast Power Control Into Capacity*

The maximum number of simultaneous users N on one carrier in a cellular CDMA system can be calculated as in [23]:

$$N = F\left[G_p\left(\frac{E_b}{N_0}\right)^{-1} + 1\right] \tag{7.16}$$

where G_p is the processing gain, E_b/N_0 is the received bit energy per noise density, and

$$F = \frac{I_{intra}}{I_{intra} + I_{inter}} \tag{7.17}$$

where I_{intra} is the intracell interference, I_{inter} is the intercell interference in a nonfading channel, and F denotes the percentage of interference originating from intracell mobiles at the base station. In this definition of F, intracell interference I_{intra} at the base station includes all the interference from the mobiles that are connected to that base station. Therefore, users in soft handover are included in intracell interference in 2 to 3 base stations.

Let this F be valid in a nonfading channel. In a fading channel the average transmission power increases, and thus, the interference from other cells I'_{inter} is higher than the intercell interference in a nonfading channel I_{inter}. Now the percentage of interference from the own cell F' is

$$F' = \frac{I_{intra}}{I'_{intra} + I_{inter}} \tag{7.18}$$

where I'_{inter} is the intercell interference in a fading channel.

The increase in interference is equal to the increase in the average mobile transmission power ΔP_{tx} shown in Table 7.6, where ΔP_{tx} is given by

$$\frac{I'_{inter}}{I_{inter}} = \Delta P_{tx} \tag{7.19}$$

Now, we can obtain F'

$$F' = \frac{1}{1 + \Delta P_{tx}\left(\frac{1}{F} - 1\right)} \tag{7.20}$$

F is here assumed to be 55% in a nonfading channel with hexagonal base stations. This value is obtained from the system simulator for the macrocell environment and is a typical value for this environment. The values for F' in different channels are shown in

Table 7.7. If we assume in (7.16) that the processing gain G_p and the E_b/N_0 requirement do not change between different channels, then $N = F\cdot$constant, (i.e., the relative capacity):

$$\frac{N'}{N} = \frac{F'}{F} \tag{7.21}$$

The relative capacities N'/N are shown in Table 7.7.

Table 7.7
Relative Capacities Compared to a Nonfading AWGN Channel With Hard Handovers

Channel model	Percentage of interference from intracell mobiles F'	Relative capacities N'/N compared to AWGN (%)
AWGN channel	55	100
1-tap Rayleigh	11	19
2-tap Rayleigh	38	69
4-tap Rayleigh	48	87
ATDMA macro	41	74
ATDMA macro with antenna diversity	48	86

7.4.2.3 *Simulated Cellular Capacity in Fading Channel*

In this subsection the analytical capacity estimates are compared to the simulated values. The power control dynamic range is 80 dB, which is larger than the value used in the analytical calculations. This is due to the additional distance-dependent attenuation as well as to the slow fading that must be compensated for in the simulator by the available power control dynamic range. The parameters of the power control scheme are shown in Table 7.8.

The propagation model of the system simulator consists of attenuation, shadowing, and statistically generated fast fading. The attenuation model used is presented in Section 7.3.2.

Even if the power control algorithm in the simulations has a perfect knowledge of the received SIR, the fast fading process cannot be followed perfectly because of a finite step size.

Table 7.8
Simulation Parameters for Fast Power Control Simulations

Power control frequency	1 kHz
Power control step size	1.0 dB
Power control dynamics	80 dB
Mobile speed	1 km/h
Active set size in soft handover	3
Handover margin in non-ideal handover	5 dB
Log normal shadowing	mean 0 dB, std dev 6 dB
User bit rate	10 Kbps

Both hard and soft handovers are supported by the system simulation. When soft handover is used, the mobile station is able to communicate simultaneously with one or more base stations. The uplink transmission power in the mobile station is increased only if all base stations in the active set request more transmission power through fast power control signaling. Uplink macro diversity in soft handover is considered by selecting the best source (the frame with the highest average SIR) on a frame-by-frame basis. Each active base station receives the frame transmitted by the mobile station. The quality of the frame is measured by calculating the average SIR over the frame. The frame with the highest quality is utilized, while other frames are discarded.

Hard handover is modeled in simulations by connecting the mobile station to the base station to which the distance dependent attenuation and the slow fading attenuation is the lowest. This can be done by setting the maximum active set size to one.

Simulated and calculated capacities for hard handovers can be compared in Table 7.9. The simulated capacities are lower than the calculated capacities except in the 1-tap Rayleigh fading channel. These differences are due to nonidealities in the power control in the network simulation. The power control step size was fixed and the same in different fading channels. It should be noted that the SIR target was set to be the same for all capacity simulations. In practice, however, the 1-tap Rayleigh type channel with power control imperfections would require the highest SIR target for power control to achieve comparable link quality with other cases. In other channels, the power control step size was too large, and thus, unnecessary power fluctuations increased the interference and resulted in lower capacities.

Soft handover (macro diversity) shows a considerable gain in the 1-tap Rayleigh fading channel over hard handover.

The differences in capacities between different channel models point out the importance of including fast fading, fast power control, and realistic handovers in CDMA system simulations if realistic capacity results are needed for network planning. The simulated capacity with 1-tap Rayleigh fading channel and soft handover is only 43% of that calculated and simulated for nonfading channel even if the same SIR target was used in the system simulations.

The results of Table 7.9 and Figure 7.8 show capacity differences between different amounts of diversity. Therefore, if the channel does not have enough diversity, other means such as receiver antenna diversity or transmit diversity need to be used. Furthermore, in order to obtain reliable capacity results the multipath channel needs to be modeled in the same way both in link level simulations as well as in system level simulations.

Table 7.9
Simulated Relative Capacities With Different Channel Models

Channel model	Hard handover (slow) (%)	Soft handover (fast) (%)
Non-fading AWGN channel	100	100
1-tap Rayleigh	25	43
2-tap Rayleigh	57	Not simulated
4-tap Rayleigh	78	Not simulated
ATDMA macro	60	Not simulated
ATDMA macro with antenna diversity	79	Not simulated

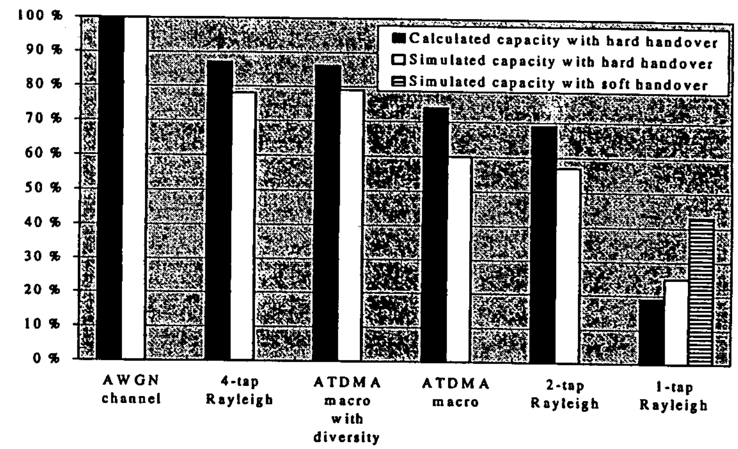


Figure 7.8 Calculated and simulated relative capacities with different channel models.

7.4.3 Capacity Degradation due to Imperfect Power Control

One purpose of power control is to equalize the received powers in the base station to eliminate the near-far effect. However, in practice it cannot eliminate the near-far effect completely due to imperfections in the power control loop. Thus, the residual variation in SIR causes degradation of the system capacity [24]. The amount of residual variation, and thus capacity degradations, depends on channel response, the system chip rate, SIR estimation accuracy, power control loop delay, command rate, and power control step size. The effect of imperfect power control is included in the E_b/N_0 results of the link level simulation since a realistic power control algorithm has been used.

The effect of imperfect power control has been studied widely in connection with IS-95 uplink performance analysis. As indicated in several studies, the performance in terms of system capacity decreases by the percentages listed in Table 7.10, as the power control error increases. The power control error causes variations in the received E_b/N_0. Empirical evidence suggests that E_b/N_0 is log-normally distributed and that the standard deviation for the narrowband CDMA employing both open loop and fast closed loop power control is between 1.5 and 2.5 dB [25].

Wideband CDMA reduces the impact of imperfect power control. This is because better diversity results in less variation in residual SIR. In [24] power control feedback delay was shown to be less critical for wideband systems than narrowband systems.

Table 7.10
Capacity Reduction With Power Control Error

Power control error, standard deviation (dB)	Capacity reduction (%)
0	0
1	3
1.5	8
2.0	13
2.5	20

Source: [26]

MUD alleviates the near-far effect, and thus the degradation of performance due to imperfect power control is smaller. In practice the gain of MUD against power control errors depends on the performance of MUD (i.e., how much of the generated intracell interference a particular receiver solution is able to remove).

7.5 SPECTRUM EFFICIENCY

In this section the evaluation of W-CDMA spectrum efficiency is analyzed. First, the link level performance of the uplink is simulated, taking into account multiuser detection efficiency. Data rates of 12 Kbps, 144 Kbps and 2 Mbps are considered. Second, spectrum efficiency results are presented based on the system level simulation using the link level results as depicted in Figure 7.3. Third, to verify the simulation results, analytical spectrum efficiency calculations are presented.

7.5.1 Link Level Performance of W-CDMA Uplink

As an input for the uplink capacity studies, link level performance, E_b/N_0 for the desired BER, and MUD efficiency are needed. The performance was studied in both CODIT macro and microcell environments. The performance figures are given in Figures 7.9 and 7.10.

The BER performances in COSSAP Monte Carlo simulations are used to assess the efficiency of multiuser detection β. This efficiency denotes the percentage of intracell interference being removed by multiuser detection at the base station receiver. The efficiency of MUD is estimated from the load that can be accommodated with a specific E_b/N_0 value with a conventional RAKE receiver and with a multiuser receiver. The same efficiency of MUD is assumed for different loads in the system simulator. The additive white Gaussian noise N_0 is used to represent both thermal noise and intercell interference, while intracell interference is represented by real transmitters. The target BER is 10^{-3}. In the analysis, the number of users with a RAKE receiver is denoted by K_{RAKE} and that with a MUD receiver by K_{MUD}. The efficiency of MUD β at a given E_b/N_0 value is given by

$$K_{RAKE} = (1-\beta)K_{MUD} \tag{7.22}$$

This applies to the macrocell where multipath interference is significant. In a microcell, the channel model is close to a single path channel and thus, self-interference from multipath components is negligible. Therefore, we can subtract the desired user from the total number of users and calculate the MUD efficiency for the microcell environment by

$$K_{RAKE} - 1 = (1-\beta)(K_{MUD} - 1) \tag{7.23}$$

To keep the complexity of the receiver to a moderate level, only four RAKE fingers were used in the macrocell with which 63% of the total energy could be collected on average. In the microcell, two RAKE fingers were used with which 95% of the energy could be captured. If only part of the energy is captured by the receiver, a higher E_b/N_0 is required, resulting into lower capacity figures. Therefore, the required E_b/N_0 in the macrocell is higher than in microcell. In the CODIT macrocell channel, the capacity was increased by about 30% by using nine RAKE fingers instead of four fingers [27].

We performed the link level simulations for the MUD efficiency study with some differences to the assumptions mentioned earlier. The service studied in this case was 74 Kbps in the CODIT macro- and microcell environments with BPSK data and chip modulation, and antenna diversity with a 0.7 correlation factor. Pilot symbols for channel estimation were inserted in the data stream. It should be noted that since delay estimation was assumed to be perfect, these results are only indicative. Furthermore, the MUD efficiency varies according to system load, which has not been considered in these simulations. The multiuser detection algorithm that has been used in the link level simulations has been presented in [28]. The impact of imperfect delay estimation on the multiuser detection has been considered, for example, in [29,30].

Simulated BER curves for the microcell are shown in Figure 7.9 and for the macrocell in Figure 7.10. All the users have equal average power. The bit error probabilities shown are the mean values of all the users. The single user case without intracell interference is shown for comparison.

In the microcell we have used simulations to estimate the required E_b/N_0 when 15 users are transmitting in the multiuser detection system. The capacity of the conventional RAKE receiver with the same E_b/N_0 is found to be between five and six users yielding a MUD efficiency of 64% to 71%. In the macrocell, the corresponding figures are 10 users with MUD and three to four users without MUD. The efficiency of MUD in macrocell is then 60% to 70%. The link level simulation results and the resulting MUD efficiencies are summarized in Table 7.11. Since the calculated MUD efficiency is valid for the intracell interference only, it does not directly convert to the gain in the cellular capacity. The interference from all the cells in the network must be included in the study to obtain the capacity for the whole network (see (7.27)).

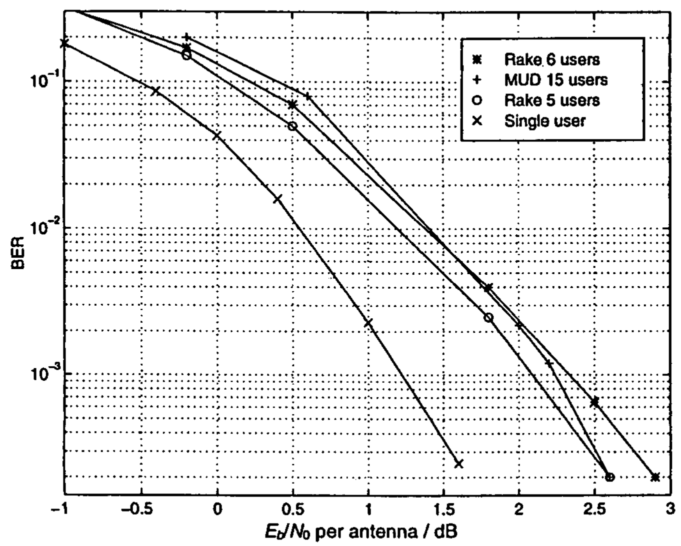


Figure 7.9 BER as a function of E_b/N_0 for the CODIT micro-cell with 36 km/h.

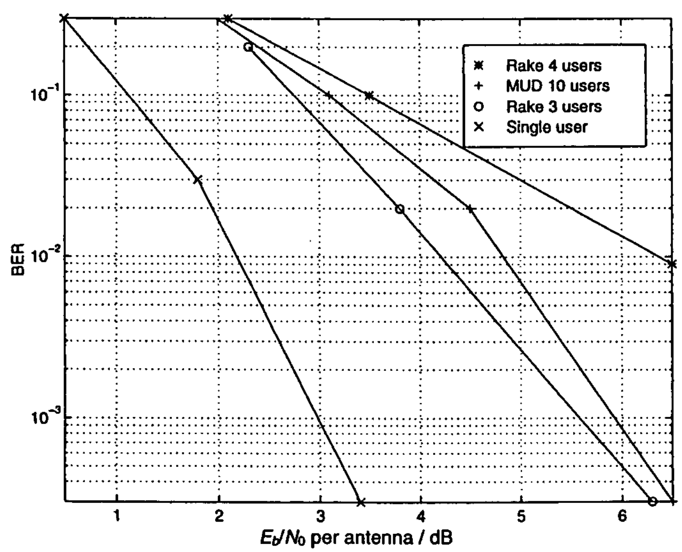


Figure 7.10 BER as a function of E_b/N_0 for the CODIT macro-cell with 50 km/h.

Table 7.11
Link Level Results With and Without Base Station Multiuser Detection

Channel model, E_b/N_0 value	Number of users, no MUD, K_{RAKE}	Number of users with MUD, K_{MUD}	Efficiency of MUD β
micro, 2.2 dB	5 – 6	15	64%-71%
macro, 5.9 dB	3 – 4	10	60%-70%

For the spectrum efficiency comparison for the uplink, results were produced for 12- and 144-Kbps services in the macrocell environment with the channel model presented in Figure 7.7. The main difference with the downlink results is the use of uplink antenna diversity. In the simulations for macrocell, the fading on the different antenna branches has been assumed to be uncorrelated. For the studied 12- and 144-Kbps services, the performance results are given in Table 7.12 for CODIT microcell multipath profile.

Table 7.12
Link Level Results of 12- and 144-Kbps in CODIT Microcell

Studied service in the uplink	E_b/N_0
12 Kbps	5.7
144 Kbps	1.7

The difference between the uplink 12- and 144-Kbps services is because of the larger interleaving period for 144 Kbps and because more energy is available for the decision directed channel estimator, thus giving better channel estimates for 144 Kbps than for 12 Kbps.

7.5.1.1 *Transmission of 2 Mbps*

Link level performance simulations for the 2-Mbps user bit rate and BER of 10^{-3} were carried out using similar wideband channel models as in the CODIT project (described in Chapter 4), for urban macrocell and urban street microcell at speeds of 3 and 36 km/h and for indoor environments [9] at a speed of 3 km/h using three different chip rates of 5, 10, and 20 Mcps. The actual target BER of 10^{-6} was not simulated because of the very high computational efforts required.

The performance of uplink with 2-Mbps service is good (4 to 6 dB without antenna diversity) with all simulated bandwidths. Receiver antenna diversity was not used in the link level simulations, but diversity gain is taken into account in the range and capacity simulations in the following sections. Fast power control offered considerable gains in the microcell and indoor environments. The E_b/N_0 performance was in the best cases within 1 dB of the simulation results in nonfading channels for the coding scheme given in [1]. Uplink simulation results for BER of 10^{-3} with fast power control and without antenna diversity are shown in Table 7.13. The corresponding results without power control are presented in Table 7.14.

Table 7.13
Uplink Performance With Fast Power Control but Without Antenna Diversity (BER 10^{-3})

Bandwidth	5.12 Mcps	10.24 Mcps	20.48 Mcps
Channel	E_b/N_0	E_b/N_0	E_b/N_0
Macro 3 km/h	4.6	5.5	6.2
Micro 3 km/h	3.9	3.6	4.1
Indoor 3 km/h	N/A	3.6	3.3

Table 7.14
Uplink Performance Without Power Control and Without Antenna Diversity (BER 10^{-3})

Bandwidth	5.12 Mcps	10.24 Mcps	20.48 Mcps
Channel	E_b/N_0	E_b/N_0	E_b/N_0
Macro 3 km/h	7.0	7.0	6.5
Micro 3 km/h	10.4 (with antenna diversity)	9.9	7.5
Indoor 3 km/h	N/A	6.0	5.3

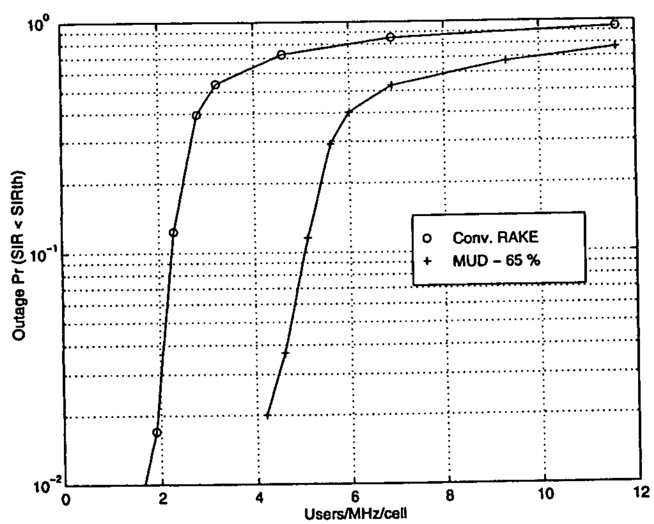


Figure 7.11 Outage probability curves for microcells with conventional receiver and with MUD receiver with $\beta = 65\%$.

7.5.2 Spectrum Efficiency of W-CDMA Uplink

Figure 7.11 shows the outage figures as a function of the load in the microcell. The capacity with a conventional RAKE receiver is 2.0 users/MHz/cell, and for the system with MUD it is 4.7 users/MHz/cell with the 74-Kbps service used in the analysis for MUD performance. The capacity is obtained with 5% outage.

In Figure 7.12 the simulated outage figures for the macrocell are presented as a function of load for different MUD efficiencies. Lines indicate performances with different MUD efficiency. It can be seen that the capacity of MUD with 65% efficiency is almost doubled compared to the conventional RAKE receiver.

The high capacity of CDMA with MUD results from the fact that most of the interference is intracell interference. Because of soft handover, users in a handover state can be included in the MUD process, which allows an extra 20% of interference to be included in the interference cancellation in the macrocellular environment. In the microcellular environment, cell separation is better, fewer users are in the soft handover state, and the amount of interference from other cells is lower than that in the macrocellular environment.

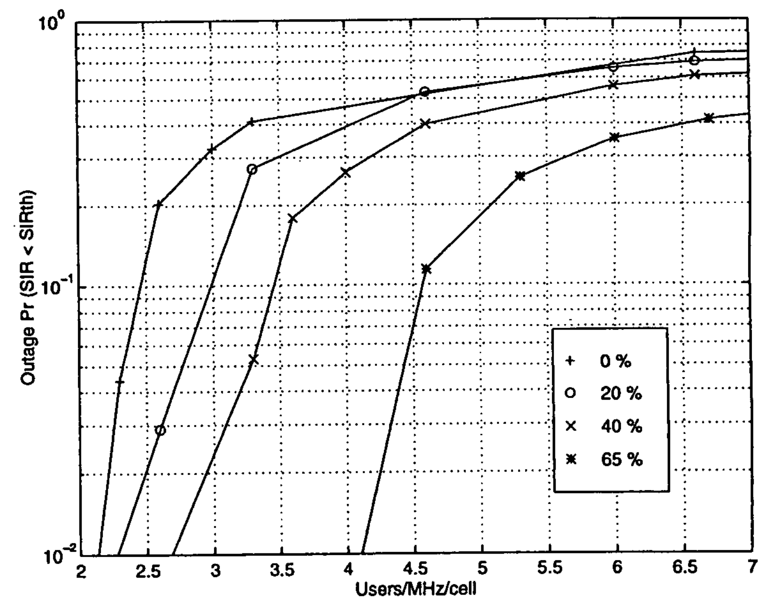


Figure 7.12 Outage probability curves as a function of MUD efficiency in macrocell. The line with 0% efficiency corresponds to the conventional RAKE receiver.

7.5.2.1 *Analytical Spectrum Efficiency Calculations*

To validate the simulation results, analytical spectrum efficiency calculations for the macrocell environment were carried out. The simulated performance compared with analytical results is given in Table 7.15. The reason for the conventional RAKE receiver having differences between simulated and analytical results is in the power control modeling. The analytical approach assumes perfect power control, whereas the simulator has the actual power control algorithm running with limited dynamics and accuracy. Furthermore, the analytical results assume a nonfading channel, while the simulated capacities are obtained in a fading channel.

Table 7.15
Analytical and Simulated (5% Outage) Capacities in Macrocell (users/MHz/cell)

	Conventional RAKE	MUD 20%	MUD 40%	MUD 65%
Simulated	2.3	2.7	3.3	4.5
Analytical	2.8	3.3	4.0	5.3

The Gaussian approximation is generally the basis for analytical capacity calculations. The central limit theorem states that the normalized sum of independent random variables approaches a Gaussian random variable as the number of terms (here users) increases [31]. This is highly applicable in the case of a large number of low bit rate users; but as the data rates get higher and the interference originates from only a few users, Gaussian approximation underestimates the interference: It should be noted again that a high degree of diversity allows the use of the Gaussian approximation a bit further because of the randomizing effect on the interference.

First, the capacity formula for the network without MUD in BTS is defined. The attained E_b/N_0 value is given by

$$\frac{E_b}{N_0} = \frac{S \cdot G_p}{I_{\text{intra}} + I_{\text{inter}} + N_0} \tag{7.24}$$

where S is the received signal strength, G_p is the processing gain, I_{intra} is the intracell interference, I_{inter} is the intercell interference from other cells and N_0 is the thermal noise. In the following, thermal noise N_0 is neglected and I_{intra} is equal to $(N-1)\,S$.

The fraction of the intracell interference caused by the users operating in the same cell as the studied user compared to the total interference is given as

$$F = \frac{I_{\text{intra}}}{I_{\text{intra}} + I_{\text{inter}}} \tag{7.25}$$

It should be noted that in (7.25) the intracell interference I_{intra} is equal to $N \cdot S$, instead of $(N-1) \cdot S$. The simulated value for F in macrocells is 0.73. For a single cell F is equal to 1. From the two equations above we get

$$\frac{E_b}{N_0} = \frac{G_p}{\frac{N}{F} - 1} \Leftrightarrow N = F \cdot \left(G_p \left(\tfrac{E_b}{N_0}\right)^{-1} + 1\right) \tag{7.26}$$

The value N is the number of users that are associated with the BTS. N also includes users that are connected to more than one BTS while in a soft handover state. The number of users being connected to two or three BTSs is obtained from the simulator and used to adjust the analysis correspondingly. Simulations show that typically in the simulated macrocellular environment, 80% of the users are connected to only one BTS, while 15% of the users are connected to two BTSs, and 5% are connected to three BTSs. The calculated capacity must be then scaled by 1.25, as the effect of soft handover is seen in F and the total number of connections in the system is higher than the number of mobiles in the system.

The corresponding analysis for the MUD case can be performed by

$$\frac{E_b}{N_0} = \frac{S \cdot G_p}{(1-\beta) I_{\text{intra}} + I_{\text{inter}} + N_0} \tag{7.27}$$

Intercell interference can be calculated by

$$I_{\text{inter}} = \frac{1-F}{F} I_{\text{intra}} \tag{7.28}$$

From (7.23) and (7.24) we have

$$\frac{E_b}{N_0} = \frac{S \cdot G_p}{(1-\beta)(I_{\text{intra}} - S) + \frac{1-F}{F} I_{\text{intra}}} \tag{7.29}$$

Since I_{intra} is equal to NS, we can write

$$\frac{E_b}{N_0} = \frac{S \cdot G_p}{(1-\beta)(NS - S) + \frac{1-F}{F} NS} \tag{7.30}$$

The capacity of the system is now given by

$$N = F \left(\frac{G_p \left(\frac{E_b}{N_0}\right)^{-1} - (\beta - 1)}{1 - F\beta} \right) \tag{7.31}$$

If β is set to 0, then the equation above becomes the same as (7.16). The β with value 0 represents the capacity of the conventional RAKE receiver-based system.

Capacity as a function of MUD efficiency is shown in Table 7.15. Both analytical and simulated results are shown. The simulated results compare well to the analytical results. The offset between analytical and simulated results is due to the non-ideal power control and due to real interferers in the simulator instead of Gaussian approximation.

Capacity strongly depends on the radio environment, which is defined by the pathloss attenuation factor, shadowing, and the wideband channel model. The proportion of the interference coming from the own cell from the total interference F in (7.31) is greatly dependent on the environment. It ranges from the single isolated cell case with no intercell interference to the macro cell environment with up to 40% of interference having intercell origin.

Multiuser detection is shown to have the potential to increase capacity. It is more suited for the uplink, where all users need to be detected in any case, and the base station has more processing power available than the mobile station. A suboptimal multiuser receiver has been shown to offer a clearly improved link performance over the RAKE receiver by removing 60% to 70% of the intracell interference in urban micro- and macrocell environments. This interference reduction achieved in the link level with a multiuser detector was used as an input to the system level simulator, which led to a considerable increase in cellular capacity. The capacity gain depends on the ratio of intracell interference to intercell interference, and therefore, the microcellular environment offering high cell isolation gained even more from the use of MUD. Alternative solutions for capacity enhancements are, for example, the use of adaptive antennas or cell splitting, both of which require extra hardware such as antennas. With the use of MUD for capacity enhancement, only the baseband hardware needs to be modified.

7.5.3 Link Level Performance of W-CDMA Downlink

As an input for the downlink system level simulations, we need link level performance results and the orthogonality factor. We first derive the orthogonality factor and then present the link level simulation results.

7.5.3.1 *Derivation of the Orthogonality Factor*

The interference in the downlink from the same BS originates from a single point and the parallel code channels can be synchronized. When using orthogonal spreading codes in the ideal case, the intracell interference could be completely avoided; but after the multipath channel, part of the orthogonality is lost and intracell interference exists also in the downlink.

For system level studies in the downlink, an estimate of the achievable degree of orthogonality needs to be drawn from the link level studies. An analytical capacity calculation of the orthogonality factor assuming Gaussian approximation is presented below. The attained SIR can be written as

$$\mathrm{SIR} = \frac{\sum_{i=0}^{N} g_i L_{p,l} P_{tx}}{\sum_{i=0}^{N} \bar{g}_i} \frac{1}{I_{total}} \tag{7.32}$$

where N is the number of perceived paths in the channel model with the selected bandwidth, $L_{p,l}$ is the pathloss between mobile stations and base station l, P_{tx} is the transmission power for the selected user, I_{total} is the total interference experienced by that user, g_i is the instantaneous path gain, and $\bar{g}_i$ the average gain for path i.

Instantaneous SIR is calculated by dividing the received signal by the interference, and multiplying by the processing gain. The downlink intracell interference is multiplied by $1-\alpha$ where α is the orthogonality factor. An orthogonality factor of 1 corresponds to perfectly orthogonal intracell users, while with the value of 0 the intracell interference is completely asynchronous. Signal-to-interference ratio is then given by

$$\mathrm{SIR} = \frac{\sum_{i=0}^{N} g_i L_{p,l} P_{tx}}{\sum_{i=0}^{N} \bar{g}_i} \frac{G_p}{(1-\alpha)I_{intra} + I_{inter} + N_0} \tag{7.33}$$

where N_0 is the thermal noise, G_p is the processing gain, I_{intra} is the intracell interference, and I_{inter} is the intercell interference. Since the system is interference limited, thermal noise N_0 is assumed small and therefore neglected. I_{intra} and I_{inter} are equal to

$$I = \sum_{l=0}^{M} \left[L_{p,l} \frac{\sum_{i=0}^{N} g_{i,l}}{\sum_{i=0}^{N} \bar{g}_i} \left(\sum_{k=0}^{R} P_{tx,k} + P_{pilot,l} \right) \right] \tag{7.34}$$

where R is the number of interferers within one cell, and $P_{tx,k}$ is the transmission power to the kth user. Each path of the desired user experiences the same interference on average. M is the number of base stations causing inter- or intracell interference and $P_{pilot,l}$ is the pilot power of base station l.

Orthogonality factor α is given by

$$\alpha = 1 - \frac{E_b}{I_0} \left(\frac{E_b}{N_0} \right)^{-1} \tag{7.35}$$

where N_0 is intercell interference, I_0 is intracell interference, and E_b/N_0 and E_b/I_0 are given in absolute figures, not in dB. E_b/N_0 is the performance figure with Gaussian noise, and E_b/I_0 is the corresponding figure dominating intracell interference. The values of E_b/N_0 and E_b/I_0 are obtained from link level simulations. The block diagram of downlink simulation with intracell and intercell interference is shown in Figure 7.13.

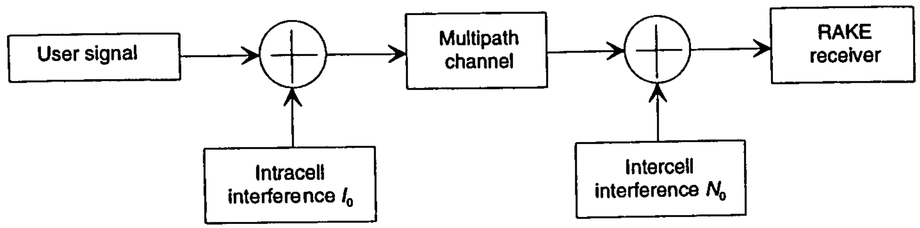


Figure 7.13 The block diagram of downlink simulation with intracell and intercell interference.

In the uplink, the signals are always asynchronous and do not have orthogonality unless the mobiles are synchronized with very high accuracy and operated in an environment where only one multipath component existed.

Figure 7.14 and Tables 7.16 and 7.17 show the results for downlink 12- and 144-Kbps link level E_b/N_0 and E_b/I_0 (in dB) in the CODIT microcell multipath channel. For 12 and for 144 Kbps results, the main differences to the uplink were the lack of antenna diversity and the use of a pilot channel for channel estimation. The pilot channel is not included in the E_b/N_0, and thus must be included in the system level modeling.

E_b/I_0 values have been calculated using the following equation:

$$E_b / I_0 \equiv \frac{W P_{user}}{R(NP + P_{pilot})} \tag{7.36}$$

where, W is the chip rate of the downlink, R is the user data rate, P_{user} is the power of the wanted signal, N is the number of interfering channels, P is the power of interfering channel, and P_{pilot} the power of the pilot signal.

Table 7.16
Downlink 12 Kbps Link Level E_b/N_0 and E_b/I_0 Results in CODIT Microcell Multipath Channel

12 Kbps E_b/I_0	12 Kbps E_b/N_0	Orthogonality factor α
2.3	6.4 dB	0.61

Table 7.17
Downlink 144 Kbps Link Level E_b/N_0 and E_b/I_0 Results in CODIT Microcell Multipath Channel

144 Kbps E_b/I_0	144 Kbps E_b/N_0	Orthogonality factor α
1.9 dB	6.1 dB	0.62

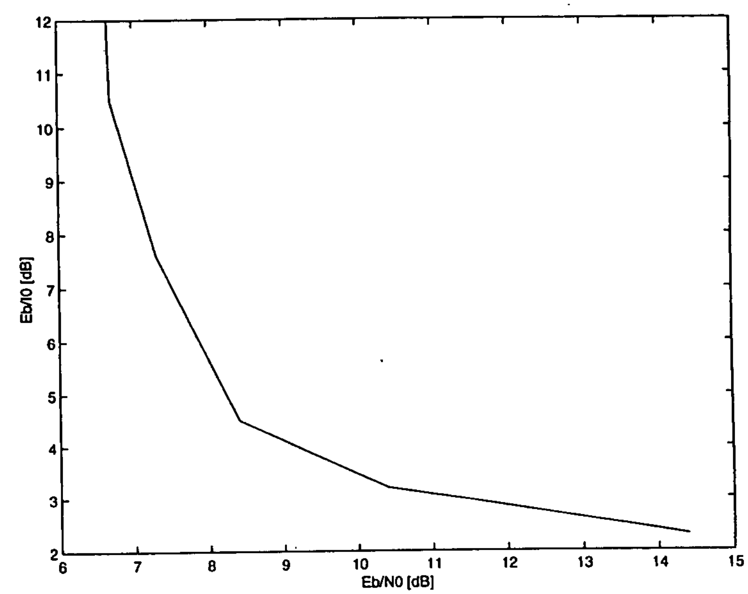


Figure 7.14 Downlink 12 Kbps with intracell interference (I_0) and with intercell interference (N_0).

7.5.3.2 *Transmission of 2 Mbps*

Transmission of 2-Mbps service was simulated in the CODIT macro-, micro-, and picocell channels. In each channel, enough RAKE fingers were allocated to gather all the signal energy. The maximum number of RAKE fingers was nine. The most important parameter used in the comparison of different configurations in noise limited cases is E_b/N_0, which represents the received energy per bit versus noise power density. In the intracell interference case, energy per bit versus intracell interference power E_b/I_0 was compared.

The corresponding link performance results are presented in Tables 7.18 and 7.19. According to the downlink simulation results, the requirements of providing 2-Mbps service are achieved using reasonable E_b/N_0 values (4 to 10 dB), even with the narrowest 5-Mcps bandwidth. Only the performance in the macro channel seems to be

somewhat poorer when the 5-Mcps bandwidth is used. This is due to severe interpath interference and low processing gain.

Table 7.18
Downlink Performance Without Power Control (BER 10^{-3})

Bandwidth	5.12 Mcps		10.24 Mcps		20.48 Mcps	
Channel	E_b/N_0	E_b/I_0	E_b/N_0	E_b/I_0	E_b/N_0	E_b/I_0
Macro 3 km/h	10.3	9.8	6.8	5.7	5.3	4.2
Macro 36 km/h	9.5	8.8	6.3	5.5	4.8	4.2
Micro 3 km/h	14.3	3.5	12.3	2.2	9.0	2.5
Micro 36 km/h	9.8	2.2	8.8	1.9	6.8	2.5
Indoor 3 km/h	9.8	0.4	9.3	-0.2	6.8	1.7

Table 7.19
Downlink Performance With Fast Power Control (BER 10^{-3})

Bandwidth	5.12 Mcps		10.24 Mcps		20.48 Mcps	
Channel	E_b/N_0	E_b/I_0	E_b/N_0	E_b/I_0	E_b/N_0	E_b/I_0
Macro 3 km/h	10.3	9.8	6.0	5.7	5.3	4.2
Macro 36 km/h	9.5	8.8	6.3	5.5	4.8	4.2
Micro 3 km/h	8.3	3.5	6.4	2.2	4.0	2.5
Micro 36 km/h	6.0	2.2	4.5	1.9	4.0	2.5
Indoor 3 km/h	4.7	0.4	4.2	-0.2	3.6	1.7

7.5.4 Spectrum Efficiency of W-CDMA Downlink

Figure 7.15 shows the downlink spectrum efficiency of 144 Kbps for micro- and macrocells as a function of the orthogonality factor. The actual orthogonality factor was calculated to be 0.7 and 0.2 in micro- and macrocells, respectively. Therefore, multiuser detection used to remove intracell interference would still improve the spectrum efficiency, since the orthogonality is not perfect (=1). This contrast often prompted results that downlink multiuser detection was not beneficial because of orthogonality.

As shown in Figure 7.16, canceling the intercell interference clearly has a smaller effect on the total system capacity. Curves indicate the capacity improvement with a different number of intercell interferers that were included in the interference cancellation process. The curves are shown as a function of the efficiency of the interference cancellation.

Link level E_b/N_0 and E_b/I_0 values used in 12- and 144-Kbps services for spectrum efficiency comparison were shown in Tables 7.16 and 7.17.

It was noted that in the downlink macrocell environment, intracell interference is most important due to multipath propagation. In the macrocell environment, significant capacity improvements could be expected if intracell interference could be canceled in the mobile. Also, in a microcell environment, intracell interference cancellation could improve the capacity but not as much as in the macrocell environment. The difference is because there is more intracell interference in macrocell due to multipath propagation.

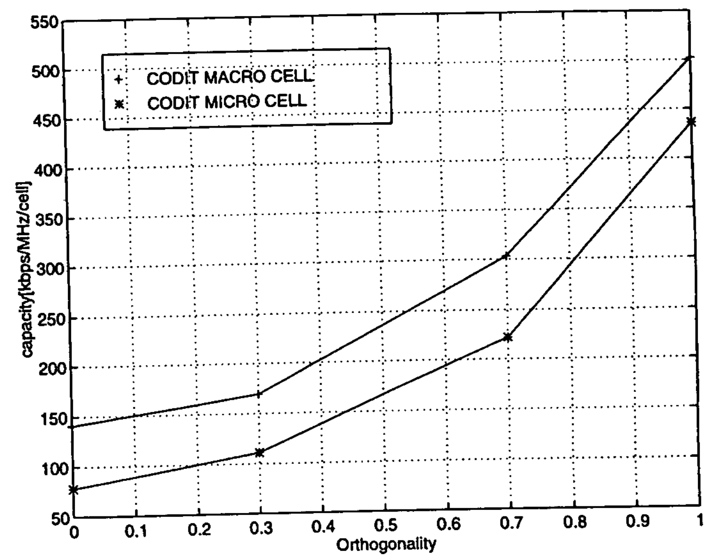


Figure 7.15 Downlink capacity as a function of intracell interference.

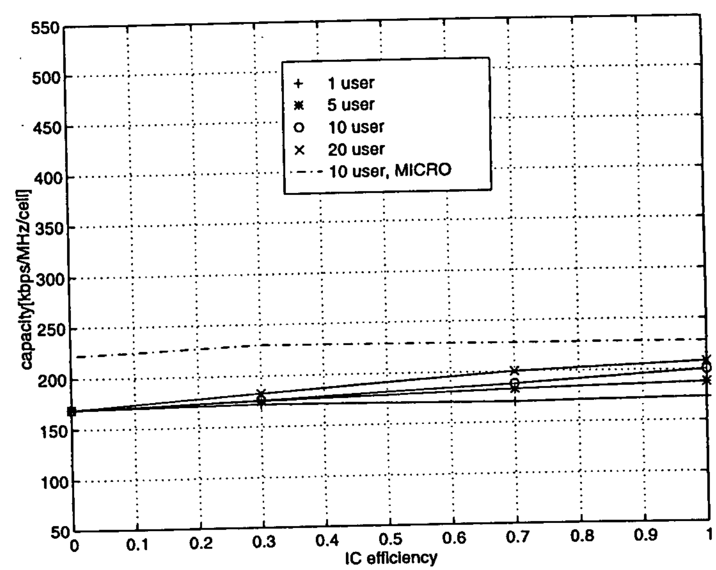


Figure 7.16 Downlink capacity as a function of intercell interference.

Additional diversity obtained from macro diversity was important in the microcell environment as there were limited number of resolvable multipath components due to large coherence bandwidth; similarly the mobile antenna diversity would be beneficial if allowed by complexity.

The downlink spectrum efficiency results were almost the same for the 12- and 144-Kbps services. This is reasonable since the downlink link level results were also very near to each other.

7.5.5 Comparison of Spectrum Efficiency for Uplink and Downlink

Table 7.20 shows capacity results for the uplink and the downlink. Downlink capacity is approximately 2 to 2.5 times lower than the corresponding uplink capacity, if both services are compared.

In addition to the base station antenna diversity, the major reason for the higher capacity of the uplink was the use of MUD, providing almost two-fold capacity if compared to a conventional receiver. In the downlink, the corresponding factor for MUD gain is the orthogonality factor. There is a difference between uplink MUD gain and downlink orthogonality in the case of soft handover. In soft handover one uplink transmission is received at two base stations, and that interference can be canceled at both base stations and MUD gain can be obtained. In the downlink there are transmissions from two base stations to one mobile station, and those two signals are not orthogonal. It should also be noticed that the orthogonality factor gives only a modest capacity gain in a macrocell channel because of multipath propagation.

Table 7.20
Cellular Capacity in Kbps/MHz/cell

	Uplink	Downlink
Macro 12 Kbps	192	108
Macro 144 Kbps	388	108

7.5.5.1 *Spectrum Efficiency With Circuit Switched 2 Mbps*

Table 7.21 shows the spectral efficiency in Kbps/MHz/cell for 2-Mbps communication for the downlink and for the uplink.

Table 7.21
Spectrum Efficiency in Kbps/MHz/cell.

	Uplink	Downlink
Micro 5 Mcps	1040	80
Micro 10 Mcps	1080	160
Macro 5 Mcps	360	40
Macro 10 Mcps	260	100

The cellular capacity of networks offering 2-Mbps circuit switched services seems to be quite low due to the poor downlink capacity. According to the results in Table 7.21, 15- to 20-MHz bandwidth is required to support one active 2-Mbps user in every cell. However, since all the users seldom transmit simultaneously, packet reservation multiple access techniques can be exploited to alleviate the capacity problem.

The difference in spectrum efficiency between uplink and downlink is due to the following features. As in the uplink, BTS antenna diversity can be used with less complexity considerations. This easily brings a 3- or 4-dB gain to the uplink direction. This difference becomes even larger if multiuser detection is applied in the base station receiver. Even in the case öf a single user, multiuser detection can be used to remove the interpath interference in multipath channels. Interpath interference is significant with low processing gain (i.e., with high bit rates). The large dynamics allowed for uplink power control also make it possible to largely compensate for the fading, especially in the microchannel case with 2 Mbps, which allows the link level results to approach the AWGN performance. In CODIT microcell, delay spread is rather narrow. As a result, the downlink receiver cannot obtain enough diversity, although the code channels remain more orthogonal. In the uplink, antenna diversity offers the required diversity.

7.6 COVERAGE

In this section, we first explain range calculation principles. Second, we consider the coverage of a DS-CDMA cellular system in both an unloaded and a loaded case. In addition, we show the impact of multiuser detection on the uplink range. If the gain from the multiuser detection is not used for coverage improvement, it can be used to reduce the mobile station transmission power. Finally, we discuss the effect of the user bit rate on the range.

7.6.1 Range Calculation

This section lists and explains the range calculation principles. The absolute range values given in this section depend on the used propagation model and cannot be directly used for network dimensioning. The values given here are relative and are to be used for comparison purposes between different services and receiver solutions. Table 7.22, which follows these principles, shows an example calculation. The parameters used in the range calculations are shown in Table 7.23 and are listed below. The link budget model is based on [3,6]. Link budget assessment is also discussed in Chapter 11.

(a) **Average transmitter power per traffic channel, $P_{TX,avg}$ (dBm).** The average transmitter power per traffic channel is defined as the mean of the total transmitted power over an entire transmission cycle with maximum transmitted power when transmitting.

(b) **Cable, connector, and combiner losses at the transmitter, L_{TX} (dB).** These are the combined losses of all transmission system components between the transmitter output and the antenna input (all losses in positive dB values).

(c) **Transmitter antenna gain, G_{TX} (dBi).** Transmitter antenna gain is the maximum gain of the transmitter antenna in the horizontal plane (specified as dB relative to an isotropic radiator).

(d) **Transmitter E.I.R.P. per traffic channel, P_{RX} (dBm).** This is the summation of transmitter power output per traffic channel (dBm), transmission system losses (dB), and the transmitter antenna gain (dBi), in the direction of maximum radiation:

$$P_{RX} = P_{TX,avg} - L_{TX} + G_{TX} \tag{7.37}$$

(e) **Receiver antenna gain, G_{RX} (dBi).** Receiver antenna gain is the maximum gain of the receiver antenna in the horizontal plane (specified as dB relative to an isotropic radiator). The received power P_{RX} at the base station can be written as

$$P_{RX} = P_{TX,avg} - L_{TX} + G_{TX} + G_{RX} \tag{7.38}$$

(f) **Cable, connector, and splitter losses at the receiver, L_{RX} (dB).** These are the combined losses of all transmission system components between the receiving antenna output and the receiver input (all losses in positive dB values).

(g) **Receiver noise figure, *NF* (dB).** Receiver noise figure is the noise figure of the receiving system referenced to the receiver input.

(h) **Thermal noise density, N_0 (dBm/Hz).** Thermal noise density is defined as the noise power per hertz at the receiver input. This can be calculated from kT, where k is the Bolzman constant and T is the temperature.

(i) **Receiver interference density, I_0 (dBm/Hz).** Receiver interference density is the interference power per hertz at the receiver front end. This is the in-band interference power divided by the system bandwidth. The in-band interference power consists of both co-channel interference as well as adjacent channel interference. Thus, the receiver and transmitter spectrum masks must be taken into account. Receiver interference density for the forward link is the interference power per hertz at the mobile station receiver located at the edge of coverage, in an interior cell. In this example, an unloaded cell is assumed. The effect of interference can also be modeled under other losses calculated as presented in Chapter 11 (Figure 11.4.)

(j) **Total effective noise plus interference density (dBm/Hz).** Total effective noise plus interference density is the logarithmic sum of the receiver noise density and the receiver noise figure and the arithmetic sum of the receiver interference density.

$$(\text{Noise} + \text{Interference}) = 10\log\left(10^{((NF + N_0)/10)} + I_0\right) \tag{7.39}$$

(k) **Information rate, $10\log(R_b)$ (dB Hz).** Information rate is the channel bit rate in (dB Hz); the choice of R_b must be consistent with the E_b assumptions.

(l) **Required $E_b/(N_0+I_0)$ (dB).** The ratio between the received energy per information bit to the total effective noise and interference power density needed to satisfy the quality objectives.

The translation of the threshold error performance to $E_b/(N_0+I_0)$ performance depends on the particular multipath conditions assumed. The CDMA uplink range can be calculated based on the link level E_b/N_0 value where E_b is energy per user bit and N_0 represents thermal and receiver noise. When reaching the maximum range, the mobile is transmitting at constant full power. Therefore, this E_b/N_0 value is obtained from link level simulations without fast power control, and it is higher than the value used in capacity simulations.

(m) **Receiver sensitivity, $P_{RX,min}$ (dBm).** This is the signal level needed at the receiver input that just satisfies the required $E_b/(N_0+I_0)$. The minimum reception power $P_{RX,min}$ represents a long time average power. If real receiver sensitivity is required, an activity factor must be added to this value. This is the case if voice activity is employed. In this case, continuous reception is assumed. The sensitivity can be calculated as

$$P_{RX,min} = kT + NF + 10\log(R_b) + E_b/(N_0+I_0) \tag{7.40}$$

(n) **Handoff gain, G_{HO} (dB).** This is the gain brought by handoff to maintain specified reliability at the boundary. In case of a single base station, there is probability P_{OUT} to have outage at distance r from the base station. In case of multiple base stations, at a point at distance r from, say, two base stations, the probability for outage relative to one base station is the same, but the probability to have outage from both the base stations is smaller and the resulting P_{OUT} is thus smaller than in the single cell case. In network planning, this reduces the shadowing margin needed in multiple cell case. The difference in shadowing margins is called handover gain. Depending on how the handover is done, the gain may be different. Here the handover gain G_{HO} is assumed to be 5 dB [10].

(o) **Other gain, Gother (dB).** An additional gain may be achieved as a result of future technologies. For instance, space diversity multiple access (SDMA) may provide an excess antenna gain.

(p) **Log-normal fade margin, ξ (dB).** The log-normal fade margin is defined at the cell boundary for isolated cells. This is the margin required to provide a specified coverage availability over the individual cells. In practice the propagation conditions vary considerably, and the pathloss attenuation factor ranges from 3 to 4 [5]. The coverage requirement of 95% sets the actual shadowing margin. If 95% coverage is to be achieved over the whole cell area and we know that we have shadowing with a standard deviation σ of 6.0 dB and pathloss model with $n = 3.6$ we get from [4] that the coverage probability at the area boundary is 84%. To have this, we need a shadowing margin of approximately 6.0 dB.

(q) **Maximum path loss, L_{max} (dB).** This is the maximum loss that permits the required performance at the cell boundary.

$$L_{max} = P_{TX} - P_{RX,min} + (G_{RX} - L_{RX}) + G_{HO} + G_{other} - \xi \tag{7.41}$$

(r) **Maximum range, d_{max} (km).** The maximum range is computed for each deployment scenario. Maximum range is given by the range associated with the maximum path loss. The propagation loss in decibels in a macrocell environment is calculated as in [32]:

$$PL = 123 + \alpha \times 10\log(d) + \sigma \tag{7.42}$$

where α is the pathloss attenuation factor, d is distance in kilometers, and σ is a Gaussian stochastic variable that models the shadow fading. It is assumed that the average of the shadow fading equals zero and the standard deviation of the shadow fading is 6 dB. However, this might vary between 6 and 10 dB [5]. The maximum range d_{MAX} is given by

$$d_{max} = 10^{\frac{L_{max} - 123 - \sigma}{36}} \tag{7.43}$$

Table 7.22
Link Budget Calculation Template

		Downlink	Uplink	Unit
(a)	Average TX power/TCH	30	24	dBm
(b)	Cable, connector, and combiner losses	2	0	dB
(c)	Transmitter antenna gain	13	0	dBi
(d)	TX EIRP/TCH = a – b + c	41.00	24.00	dBm
(e)	Receiver antenna gain	0	13	dB
(f)	Cable, connector, and splitter losses at the receiver	0	2	dB
(g)	Receiver noise figure	5	5	dB
(h)	Thermal noise density	–174	–174	dBm/Hz
(i)	Receiver interference density	–1000	–1000	dBm/Hz
(j)	Total noise + interference density (g + h + i)	–169	–169	dBm
	Information rate R_b	8	8	kHz
(k)	$10\log(R_b)$	39.03	39.03	dBHz
(l)	$E_b/(N_0 + I_0)$ (link level sim. result)	8	6.6	dB
(m)	Receiver sensitivity (j + j + l)	–122.0	–123.4	dBm
(n)	Handoff gain	5	5	dB
(o)	Other gains	0	0	dB
(p)	Log normal fade margin	11.3	11.3	dB
(q)	Maximum path loss (d – m + (e - f) + n + o)	153.67	149.07	dB
	Range	4.84	3.61	km

Table 7.23
Parameters for Range Calculations

Service	Medium bit rate data
User data rate	144 Kbps
BER	10^{-3}
Delay	100 ms
Environment	Macrocellular
Cell layout	Hexagonal
Pathloss exponent α	3.6
Log-normal shadowing σ	6 dB
Handover gain G_{HO}	5 dB
Fractional cell loading F	70%
Transceiver parameters	
Base station antenna gain G_{BS}	6 dBi
Mobile station antenna gain G_{MS}	0 dB
Maximum mobile transmission power $P_{TX,MS}$	1W = 30 dBm
Thermal noise kT	–174 dBm
Receiver noise figure NF	7 dB
Capacity with 5% outage [33]	
Downlink capacity	169 Kbps/MHz/cell = 7.0 users at 6.0 MHz
Uplink E_b/N_0 at constant (full) transmission power	4.0 dB

7.6.2 Range in Unloaded and Loaded Networks

In cellular CDMA, the achievable range is very much dependent on cell load: If the number of users increases, the range decreases. This agrees with all cellular solutions where system capacity is considered to be mainly limited by the interference generated by the system itself. With CDMA, the effect is also seen in a single cell case, unlike in systems where intracell users are completely orthogonal, like in pure TDMA-based systems.

In this section, the decrease in cell range is calculated when network load is increased, and the effect of base station multiuser detection in increasing the range in a loaded network is also analyzed. Base station multiuser detection is shown to lower the mobile station transmission power in a loaded network, thus making uplink range and average mobile transmission power more insensitive to network load. The impact of intercell interference at the base station in different propagation environments is also studied. For actual network planning, multiuser detection and propagation environment must be taken into account.

In the case of an unloaded network, the uplink direction limits the achievable range and coverage, as the maximum transmission power of the mobile station is low compared to the maximum transmission power of the base station in the downlink. In a loaded network the downlink may limit the range if there is more load and thus more interference in the downlink than in the uplink.

The received SIR at the base station is obtained by [23]

$$\frac{E_b}{I_0 + N_0} = \frac{E_b}{I_{\text{intra}} + I_{\text{inter}} + N_0} \tag{7.44}$$

where E_b is the received energy per bit, I_{intra} is the intracell interference from own cell mobiles, I_{inter} is the intercell interference from those mobiles not connected to this particular base station, and N_0 is the thermal noise.

In the case of an unloaded network, $I_{\text{intra}} = 0$, $I_{\text{inter}} = 0$, and the required E_b/N_0 for range calculations is simply equal to $E_b/(I_0 + N_0)$. In a loaded network, the percentage of own cell interference from the total interference (fractional cell loading) is defined as

$$F = \frac{I_{\text{intra}} + S}{I_{\text{intra}} + S + I_{\text{inter}}} \tag{7.45}$$

where $S = E_b/G_p$ is the received signal power from one user and G_p is the processing gain.

The value for F depends on the propagation environment. The higher the pathloss attenuation factor, the higher the F. Intracell interference I_{intra} can be given in terms of intercell interference I_{inter} as

$$I_{\text{inter}} = I_{\text{intra}}\left(\frac{1}{F} - 1\right) + \frac{E_b}{G_p}\left(\frac{1}{F} - 1\right) \tag{7.46}$$

Intracell interference is in terms of number of users N given by

$$I_{\text{intra}} = (N-1)\frac{E_b}{G_p} \tag{7.47}$$

and the total interference can be written as

$$I_{\text{intra}} + I_{\text{inter}} = \left(\frac{N}{F} - 1\right)\frac{E_b}{G_p} \tag{7.48}$$

Now, setting the required $E_b/(I_0 + N_0)$ in the loaded network equal to the required E_b/N_0 in the unloaded network gives

$$\frac{E_{b,\text{loaded}}}{N_0 + I_0} = \frac{E_{b,\text{loaded}}}{\left(\frac{N}{F} - 1\right)\frac{E_b}{G_p} + N_0} = \left(\frac{E_b}{N_0}\right)_{\text{unloaded}} \tag{7.49}$$

and solving the needed E_b/N_0 in the loaded case gives

$$\left(\frac{E_b}{N_0}\right)_{\text{loaded}} = \frac{1}{\left(\frac{E_b}{N_0}\right)_{\text{unloaded}}^{-1} - \left(\frac{N}{F} - 1\right)\frac{1}{G_p}} \tag{7.50}$$

In Figure 7.17, range is presented as a function of uplink load for fractional cell load, $F = 70\%$. Load is shown as a percentage of maximum downlink capacity.

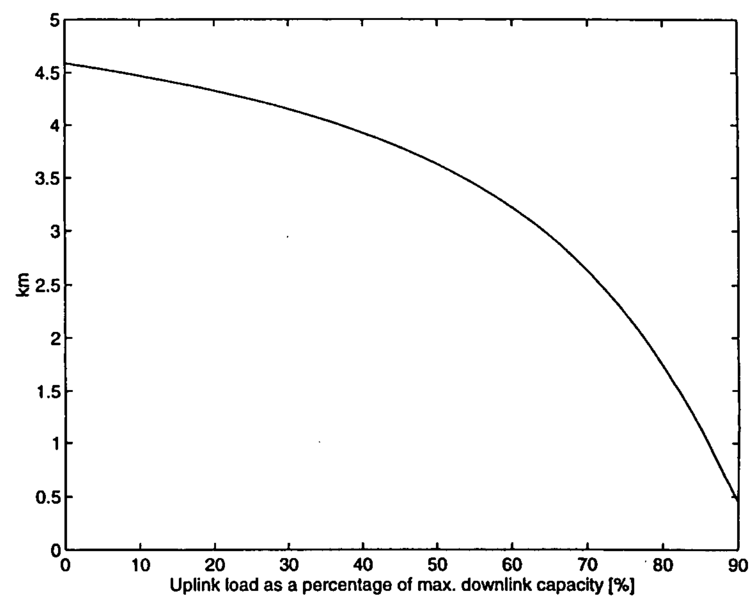


Figure 7.17 Range as a function of load. Load is shown as a percentage of maximum downlink capacity.

The maximum cell range decreases significantly when the load increases beyond about 50% of maximum downlink load. In an unloaded network the maximum cell range is 4.6 km, at 50% load range is 3.6 km, and at 90% load range is only 0.5 km. In

Figure 7.18 the E_b/N_0 requirements as a function of number of users is shown for F = 50%, 70%, and 90% corresponding to different propagation environments.

For F = 50% the increase in E_b/N_0 requirements is much higher than for F = 70% since there is, in addition to intracell interference, high intercell interference.

7.6.3 Range Extension With Base Station MUD in Loaded Networks

The range decrease as a function of load can be partially avoided by using interference cancellation or multiuser detection at the base station. Base station MUD can be used to increase the uplink capacity, but it can also be used to extend the range in a loaded network. It is assumed that MUD is able to cancel part of the intracell interference at the base station.

Base station MUD can be used to make range and cellular coverage more insensitive to uplink load. In practice, MUD efficiency β will depend on the channel estimation algorithm, interference cancellation algorithm, and mobile speed. This introduces one new parameter into network planning. Also, the actual impact of MUD into cell range depends on the fractional cell loading F, which has to be predicted for each cell individually.

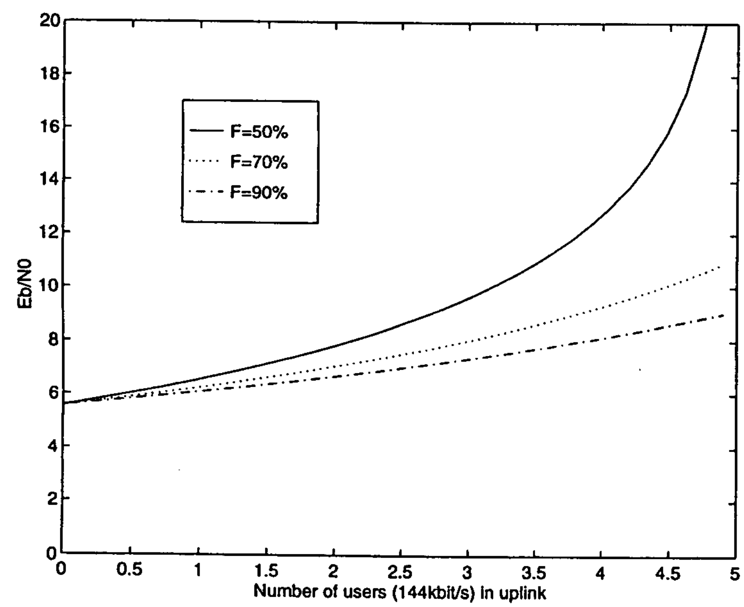


Figure 7.18 E_b/N_0 requirements in dB as a function of load in different propagation environments.

The effect of MUD can be taken into account by using the efficiency of MUD β as a measure of performance of MUD. That gives the percentage of intracell interference that is removed by MUD. With MUD the intercell interference $I_{\text{intra,MUD}}$ can be written as

$$I_{\text{intra,MUD}} = (1-\beta) I_{\text{intra}} \tag{7.51}$$

The total interference is now

$$I_{\text{intra,MUD}} + I_{\text{inter}} = \left(\frac{N(1-\beta)+\beta}{F} - 1 \right) \frac{E_b}{G_p} \tag{7.52}$$

The required E_b/N_0 in the loaded network with MUD becomes

$$\left(\frac{E_b}{N_0} \right)_{\text{loaded,MUD}} = \frac{1}{\left(\frac{E_b}{N_0} \right)_{\text{unloaded}}^{-1} - \left(\frac{N(1-\beta)+\beta}{F} - 1 \right) \frac{1}{G_p}} \tag{7.53}$$

Range as a function of efficiency of base station MUD is shown in Figure 7.19. An efficiency of 0% corresponds to the base station without MUD. We can see that the range varies considerably depending on MUD efficiency.

From Section 7.5 we obtain a simulated figure for efficiency of MUD β: 70%. Now, the ranges with and without MUD are presented in Table 7.24 and in Figure 7.20.

Table 7.24
Range With and Without Base Station MUD.

Uplink load (%)	Range without MUD (km)	Range with MUD (km) (β=70%)
0	4.6	4.6
50	3.6	4.4
70	2.6	4.2
90	0.5	4.0

7.6.4 Mobile Transmission Power Savings with Base Station MUD

Even if the required coverage in the network can be achieved without base station MUD, advanced receiver algorithms at the base station can be used to decrease the transmission power of the mobile stations. The transmission power at maximum range $P_{\text{TX,MS}}$ can be written from (7.36)

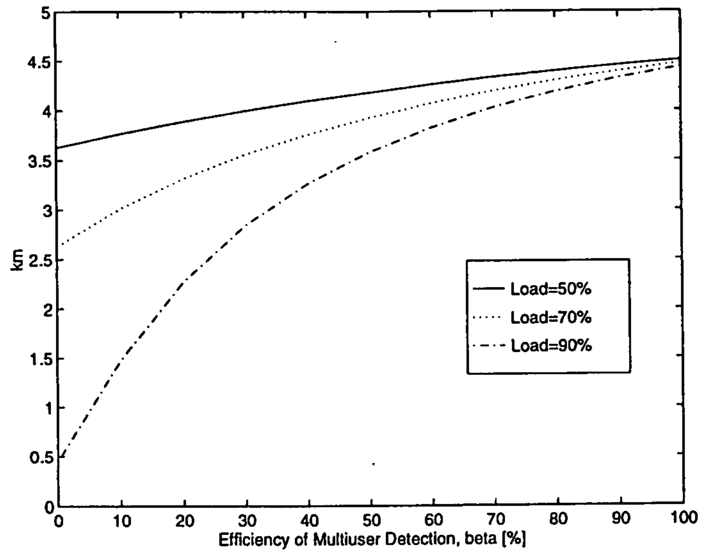


Figure 7.19 Range extension with base station MUD.

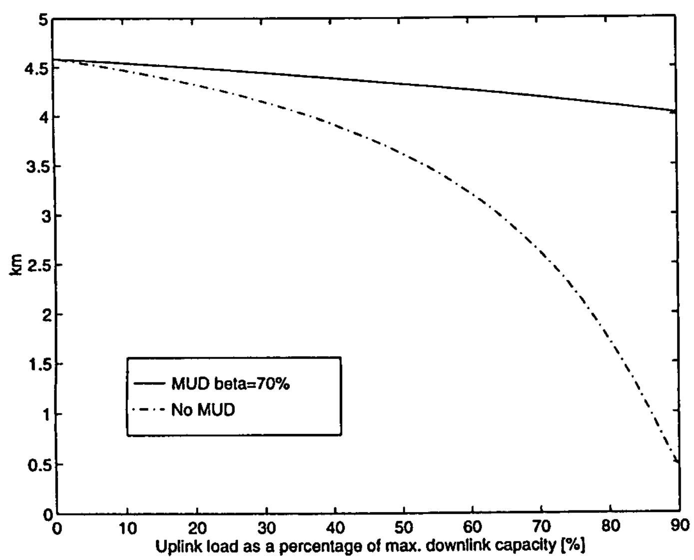


Figure 7.20 Range as a function of load with and without MUD.

$$P_{\mathrm{TX,MS}} = P_{\mathrm{RX,min}} + L_{\mathrm{HO}} - G_{\mathrm{MS}} - G_{\mathrm{BS}} \tag{7.54}$$

and by replacing $P_{\mathrm{RX,min}}$ with (7.35)

$$P_{\mathrm{TX;MS}} = \frac{E_b}{N_0} + R_b + NF + kT + L_{\mathrm{HO}} - G_{\mathrm{MS}} - G_{\mathrm{BS}} \tag{7.55}$$

Since all the terms in (7.50) are the same except for E_b/N_0, regardless of the base station receiver algorithm, $P_{\mathrm{TX,MS}}$ is determined only by the E_b/N_0 requirement. The decrease in required transmission power with MUD is therefore

$$\frac{P_{\mathrm{TX,MS}}}{P_{\mathrm{TX,MS,MUD}}} = \frac{\left(E_b/N_0\right)_{\mathrm{loaded}}}{\left(E_b/N_0\right)_{\mathrm{loaded,MUD}}} \tag{7.56}$$

and the gain in E_b/N_0 can be written as

$$\frac{\left(\frac{E_b}{N_0}\right)_{\mathrm{loaded}}}{\left(\frac{E_b}{N_0}\right)_{\mathrm{loaded,MUD}}} = \frac{\left(\frac{E_b}{N_0}\right)^{-1}_{\mathrm{unloaded}} - \left(\frac{N(1-\beta)+\beta}{F} - 1\right)\frac{1}{G_p}}{\left(\frac{E_b}{N_0}\right)^{-1}_{\mathrm{unloaded}} - \left(\frac{N}{F} - 1\right)\frac{1}{G_p}} \tag{7.57}$$

In the above equations, it must be assumed that the propagation loss is lower than the maximum loss allowed for a system without MUD. If the mobile is not transmitting at full power but is able to use fast power control, then (7.52) is not exact but an approximation. In Figure 7.21 the estimated decrease in average transmission power of mobile station is shown as a function of MUD efficiency.

Assuming MUD efficiency $\beta = 70\%$ and fractional cell loading $F = 70\%$, the decrease of transmission power is 3 dB at 50% load and 7 dB at 70 % load.

7.6.5 Effect of User Bit Rate on the Range

The maximum range depends on the user bit rate. The higher the user bit rate, the shorter the range. The goal of third generation networks is to provide full coverage for low bit rates (<144 Kbps) and for higher bit rates a limited coverage may be acceptable. In Table 7.25 the maximum ranges for different bit rates are calculated assuming the same E_b/N_0 target of 4.0 dB for all services and the parameters in Table 7.23.

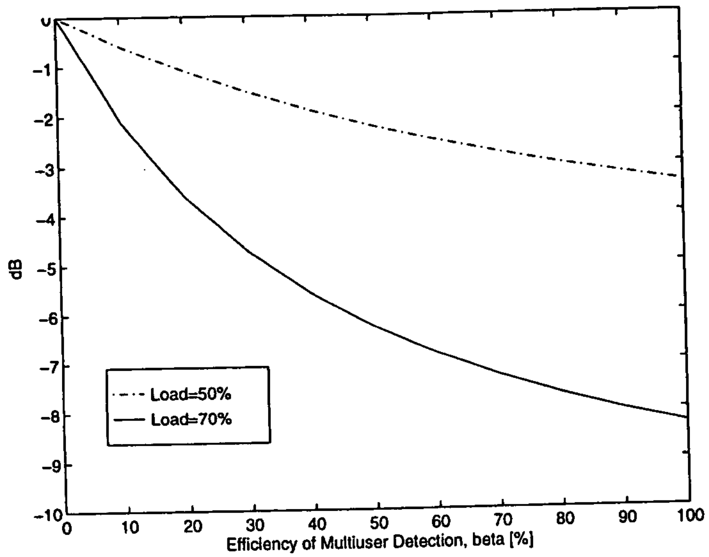


Figure 7.21 Savings in mobile transmission power with MUD.

Table 7.25
Maximum Range as a Function of User Bit Rate.

User bit rate	Relative maximum range (km)
8 Kbps	10.0
14.4 Kbps	8.5
64 Kbps	5.6
144 Kbps	4.5
384 Kbps	3.4
1 Mbps	2.6
2 Mbps	2.2

7.6.6 Summary

Coverage of a loaded DS-CDMA cellular network has been considered by analyzing the uplink range. Base station multiuser detection (MUD) has been proposed as an upgrade solution to provide good coverage even with high system load after the initial deployment. MUD has also been shown to decrease mobile station transmission power in a loaded network. Therefore, a network with base station MUD can be operated with a higher percentage of maximum load if the system capacity is limited by downlink.

Good coverage need not be sacrificed in order to use high system loads. For actual network planning we need to take into account the performance of MUD to be able to predict and plan the coverage. The impact of MUD into cell range depends also on the propagation environment (i.e., the percentage of the intracell interference from the total interference). The increase in the data rate will inevitably mean reduced range in the uplink as the transmission power is limited. Thus, in the cell design the coverage area for the low rate services is likely to be different than for high rate service. This effect has not been visible in the second generation networks since services are narrowband services with similar range performance.

REFERENCES

[1] Proakis, J. G., *Digital Communications*, 3rd ed., New York: McGraw-Hill, 1995.

[2] Xia, H. H., "Reference Models for Evaluation of Third Generation Radio Transmission Technologies," *Proceedings of ACTS Summit*, Aalborg, Denmark, October 1997, pp. 235–240.

[3] Recommendation ITU-R M.1225, "Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000," *Question ITU-R 39/8*, 1997.

[4] Jakes, W. C., *Microwave Mobile Communications*, New York: Wiley, 1974.

[5] Parsons, J., *The Mobile Radio Propagation Channel*, London: Pentech Press, 1992.

[6] ETSI/SMG-5, Selection procedures for the choice of radio transmission technologies of the UMTS, UMTS 30.03, v. 3.0.0, May 1997.

[7] Martin, G., and M. Faulkner, "1.9 GHz Measurement-based Analysis of Diversity Power versus the number of RAKE Receiver Tines at Various System Bandwidths," *Proceedings of PIMRC'97*, September 1997, Helsinki, Finland, pp. 1069–1073.

[8] Otmani, M., and M. Lecours, "Indoor Radio Impulse Response Measurements with Polarization Diversity," *Proceedings of VTC'96*, Atlanta, Georgia, USA, May 1996, pp. 151–154.

[9] CEC Deliv. No. R2020/TDE/PS/DS/P/040/b1, CODIT Final Propagation Model, June 1994.

[10] Rege, K., S. Nanda, C. F. Weaver, and W. C. Peng, "Analysis of Fade Margins for Soft and Hard Handoffs," *Proceedings of IEEE 6th Personal, Indoor and Mobile Radio Communications Conference (PIMRC'95)*, Toronto, Canada, 1995, pp. 829–834.

[11] Chopra, M., K. Rohani and J. D. Reed, "Analysis of CDMA Range Extension due to Soft Handoff," *Proceedings of VTC'95*, 1995, pp. 917–921.

[12] Andoh, H., M. Sawahashi, and F. Adachi, "Channel Estimation Using Time Multiplexed Pilot Symbols for Coherent Rake Combining for DD-CDMA Mobile Radio," *Proceedings of PIMRC'97*, Vol. 3, Helsinki, Finland, September 1997, pp. 954–958.

[13] Tiedemann, E. G., Jr., Y-C. Jou, and J. P. Odenwalder, "The Evolution of IS-95

to a Third Generation System and to the IMT-2000 Era," *Proceedings of ACTS Summit*, Aalborg, Denmark, October 1997, pp. 924–929.

[14] Hämäläinen, S., P. Slanina, M. Hartman, A. Lappeteläinen and H. Holma, "A Novel Interface Between Link and System Level Simulations," *Proceedings of ACTS Mobile Communications Summit '97*, October 7–10, 1997, Aalborg, Denmark, pp. 599–604.

[15] Hata, M., "Empirical Formula for Propagation Loss in Land Mobile Radio Services," *IEEE Transactions on Vehicular Technology*, August 1980, pp. 317–325.

[16] Ariyavisitakul, S., "SIR Based Power Control in a CDMA System," GLOBECOM'92, Orlando, Florida, USA, December 1992.

[17] Shapira J., and R. Padovani, "Spatial Topology and Dynamics in CDMA Cellular Radio," *Proceedings in VTC'92*, Denver, CO, May 1992, pp. 213–216.

[18] Special Issue on Smart Antennas, *IEEE Personal Communications*, Vol. 5, No. 1, February 1998.

[19] Ojanperä, T., K. Rikkinen, H. Hakkinen, K. Pehkonen, A. Hottinen and J. Lilleberg, "Design of a 3rd Generation Multirate CDMA System with Multiuser Detection, MUD-CDMA," *Proceedings of ISSSTA'96*, Vol. 1, Mainz, Germany, September 1996, pp. 334–338.

[20] Nikula, E., A. Toskala, E. Dahlman, L. Girard and A. Klein, "FRAMES Multiple Access for UMTS and IMT-2000," *IEEE Personal Communications Magazine*, April 1998, pp. 16–24.

[21] Jalali, A., and P. Mermelstein, "Effects of Diversity, Power Control, and Bandwidth on the Capacity of Micro Cellular CDMA System," *IEEE Journal on Selected Areas in Commun.*, Vol. 12, No. 5, June 1994, pp. 952–961.

[22] Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, 3rd ed., New York: McGraw-Hill, 1991.

[23] Hämäläinen, S., H. Holma, and A. Toskala, "Capacity Evaluation of a Cellular CDMA Uplink with Multiuser Detection," *Proceedings of ISSSTA'96*, Mainz, Germany, September 1996, pp. 339–343.

[24] Gass, J. H., D. L. Noneaker, and M. B. Pursley, "Spectral efficiency of a Power-Controlled CDMA Mobile Communication System," *IEEE Journal on Selected Areas in Commun.*, Vol. 14, No. 3, April 1996, pp. 559–569.

[25] Padovani, R., "Reverse Link Performance of IS-95 Based Cellular System," *IEEE Personal Communications*, Vol. 1, No. 3, Third Quarter, 1994, pp.28–34.

[26] Viterbi, A. M., and A. J. Viterbi, "Erlang Capacity of a Power Controlled CDMA System," *IEEE Journal on Selected Areas in Communications*, Vol. 11, No. 6, August 1993, pp. 892–899.

[27] Ojanperä, T., P. Ranta, S. Hämäläinen, and A. Lappeteläinen "Analysis of CDMA and TDMA for 3rd Generation Mobile Radio Systems," *Proceedings VTC'97*, May 1997, Phoenix, Arizona, USA, pp. 840–844.

[28] Holma, H., A. Toskala, and A. Hottinen, "Performance of CDMA Multiuser Detection with Antenna Diversity and Closed Loop Power Control," *Proceeding of VTC'96*, Atlanta, Georgia, USA, May 1996, pp. 362–366.

[29] Parkvall, S., E. Ström, and B. Ottersen, "The impact of Timing Errors on the Performance of Linear DS-CDMA Receivers," *JSAC*, Vol. 14, No. 8, October

1996, pp. 1660–1668.

[30] Buehrer, M., A. Kaul, S. Striglis, and B. D. Woerner, "Analysis of DS-CDMA Parallel Interference Cancellation with Phase and Timing Errors," *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 8, October 1996, pp. 1522–1535.

[31] Simon, M., J. Omura, R. Scholtz, and B. Levitt, *Spread Spectrum Communications Handbook*, New York: McGraw-Hill, 1994.

[32] Pizarroso, M., J. Jiménez (eds.), "Common Basis for Evaluation of ATDMA and CODIT System Concepts," *MPLA/TDE/SIG5/DS/P/001/b1*, MPLA SIG 5, September 1995.

[33] Hämäläinen, S., H. Holma, A. Toskala, and M. Laukkanen, "Analysis of CDMA Downlink Capacity Enhancements," *Proceedings of PIMRC'97*, Helsinki, Finland, September 1997, pp. 241–245.

Chapter 8

HIERARCHICAL CELL STRUCTURES

8.1 INTRODUCTION

A third generation system must be able to support a wide range of services in different radio operating environments. Different types of cells are needed for different requirements: large cells guarantee continuous coverage, while small cells are necessary to achieve good spectrum efficiency and high capacity. Small range cells are used by low mobility and high capacity terminals, while high range cells serve high mobility and low capacity terminals. In addition, different cell types should be able to operate one upon another. Hierarchical cell structure (HCS) describes a system where at least

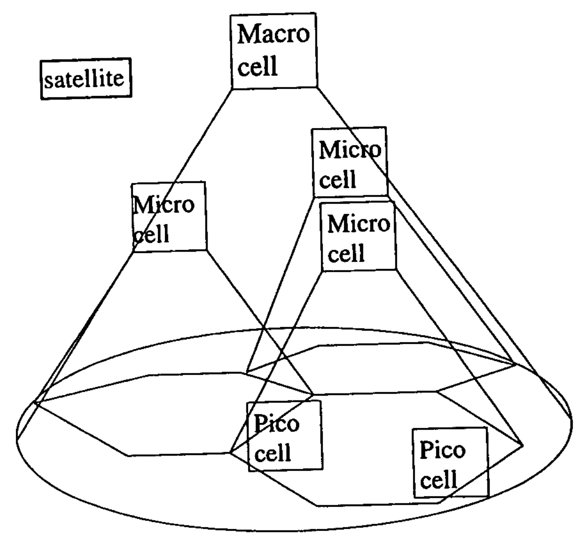


Figure 8.1 Hierarchical cell scenario.

two different cell types (e.g., macrocell and microcell) operate one upon another, as shown in Figure 8.1. Microcells are small cells covering areas with a radius of a few hundred meters. Low powered base stations are typically placed at lamp post level and they serve a street block. Macrocells, or umbrella cells, cover a radius of 1 km or more. Macrocells cover rural areas, provide continuous coverage of areas that are covered by microcells, and serve rapidly moving users. Picocells cover indoor areas with a cell radius of a few dozen meters. A low power base station usually covers an office, a floor of a high building, or a residence. Satellite cells introduce global continuous coverage by covering rural areas where terrestrial systems cannot be installed. Whenever possible, traffic should be directed to the smallest available cell, so that the system's spectral efficiency is improved. Particularly in the satellite segment, it is crucial to direct traffic to the macro- or microcells whenever possible [1].

In this chapter, two approaches for HCS design in CDMA are described, and advantages and disadvantages of the two methods are discussed. In the first approach, co-existing hierarchy layers operate in the same frequency band [2,3]. Users operating on different layers are separated by handovers and signal fading. In the second approach, different hierarchy levels are separated in the frequency domain. The aspects related to HCS are also relevant in a multioperator environment where an operator has to consider interference from the adjacent frequencies belonging to another operator. Network planning aspects such as guardbands related to multioperator environments are discussed in Chapter 11.

First, nonlinear power amplifiers are studied in Section 8.2 to establish an understanding of the adjacent channel interference mechanism in a system with different cell layers at different frequencies. A HCS system with micro- and macrocells in the same frequency, and at different frequencies, are studied in Sections 8.3 and 8.4, respectively. It should be noted that aspects related to interfrequency handover are not covered here but rather in Chapters 2 and 5. Section 8.4 evaluates link and system level performance of a cellular HCS network utilizing DS-CDMA and multiuser detection.

8.2 NONLINEAR POWER AMPLIFIERS

Even if hierarchy layers have different carrier frequencies, some adjacent channel interference (ACI) is generated between adjacent channel carriers. When bandlimited linear modulation methods are used, spectrum leakage between adjacent channel carriers depends on the linearity of the power amplifier (PA) [4]. Spectrum leakage to adjacent channels can be controlled by backing off the PA or by using some linearization method to equalize the nonlinearities of the power amplifier. Unfortunately, both methods decrease the achievable PA efficiency when compared to modulation schemes that have constant envelope and can be amplified with power efficient but nonlinear power amplifiers [4].

Real power amplifiers are always nonlinear. Linearity requirements of a CDMA transmitter are mainly determined by the spectrum ACI attenuation requirements, rather than from link level performance losses. If low ACI attenuations (i.e., larger spectrum spreading into adjacent carriers) can be tolerated, more efficient power amplifiers can be used in mobile terminals.

For link level simulations, a model of a nonlinear power amplifier is needed. In order to get real life amplifier model, a power amplifier IC was measured. The power amplifier IC is manufactured with the Gallium Arsenide heterojunction bipolar transistor (HBT) process, and the final stage of the IC is biased to class-AB.

8.2.1 Power Amplifier Characteristics

A typical power amplifier introduces both amplitude and phase distortion into the transmitted signal, resulting in AM-AM and AM-PM conversion effects. Figure 8.2 shows the AM-AM and AM-PM characteristics of the measured amplifier. The power amplifier amplifies input power linearly when input RMS power is −15 to 0 dBm. This is the linear area of the amplifier. When the input power is 0 to 5 dBm, the amplifier operates in its conversion or saturation area. The measured power amplifier has saturated output power at the level of 34 dBm. This value corresponds to 0 dB output backoff value [4]. Output backoff is defined as the difference between output power in the saturation point and output power in the operation point. Backing off the amplifier reduces output powers in the operation point and increases output backoff.

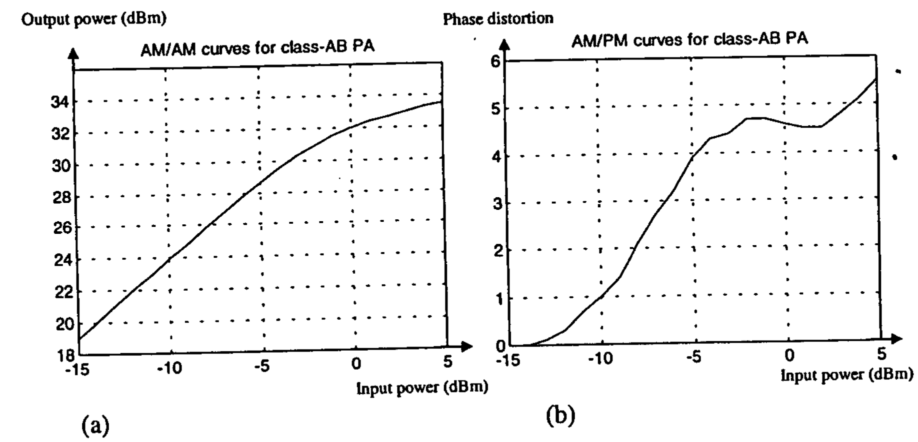


Figure 8.2 a) AM-AM curve and b) AM-PM curve for measured power amplifier IC.

8.2.2 Power Amplifier Efficiency

Efficiency of the power amplifier IC is measured using a constant envelope sine wave. Efficiency compared to output backoff is shown in Figure 8.3. At saturation power (0 dB output backoff), the power amplifier has slightly less than 50% efficiency. This means that if constant envelope modulation methods, such as Gaussian minimum shift keying (GMSK), were used, up to 50% efficiencies could be achieved. However,

because bandlimited linear modulation methods have nonconstant RF-envelope, output power needs to be lowered in order to avoid spectrum spreading [4]. The accurate efficiency of the power amplifier with bandlimited linear modulation methods (nonconstant RF-envelope) can be predicted by using the measured efficiency curve and the simulated pdfs of the used modulation method [4].

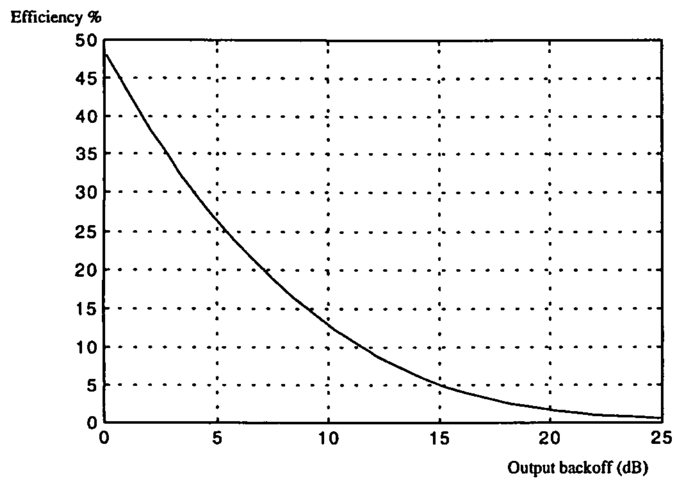


Figure 8.3 Power amplifier efficiency versus output backoff, class=AB PA.

8.3 MICRO- AND MACROCELLS AT THE SAME FREQUENCY

In the case of micro- and macrocells operating at the same frequency, the frequency reuse factor is set to one and spatial isolation is used for separating micro- and macrocell layers [2,3]. A reuse factor of one can be offered since processing gain allows users to experience interference originating from any cell layer. Intra-layer interference is controlled by power control and inter-layer interference by spatial isolation. In a generic flat-Earth model, the average transmission loss follows R^{-2} until a breakpoint that marks the separation between two segments. After the breakpoint, R^{-4} is followed. The location of the breakpoint depends on receiver and transmitter heights. As the base station is installed lower, the breakpoint occurs nearer the base station. Thus, the signal from a microcell base station attenuates faster than the signal from a macrocell base station. In Figure 8.4, spatial isolation is depicted for both hotspot case and for continuous coverage of microcells.

The problem with the presented method is that the mobile station should be instantly handed to the corresponding cell layer when it arrives to the intersection of micro- and macrocell attenuation curves. Soft handover eases the problem, but if the mobile station arrives in the microcell area, it is most often connected to both micro and

macrocell base stations. This occurs since the handover region has to be rather large in order to avoid ping-ponging between cells (or cell layers), especially if the mobile stations are fast moving. Thus, the area where mobile stations are in a soft handover state is rather wide compared to the area of microcell.

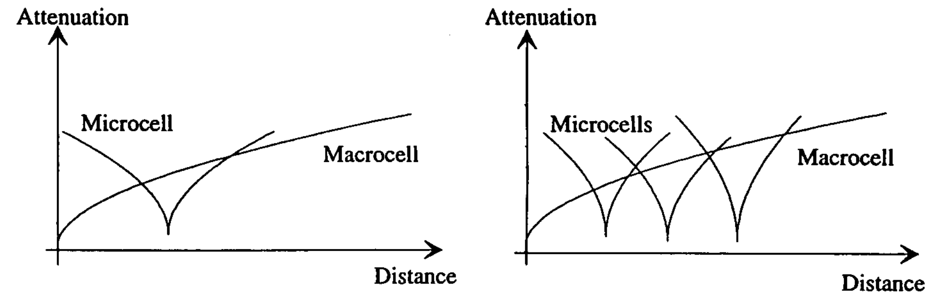


Figure 8.4 Spatial separation of umbrella cells and underlay microcells along [2]. The figure on the left depicts hotspot case covered by one microcell base stations, and the figure on the right depicts the full coverage of microcells.

Handover can never be very fast. If the number of candidate base stations is high, the pilot signal of a candidate base station can be measured only after some delay. In addition, this pilot has to be measured for a rather long period in order to filter fast fading. Handover can not be performed when the mobile station encounters a fading dip due to fast fading. After handover decision has been taken, signaling between the mobile station and the base station takes place. Before both the network side and the mobile side get all the handover related parameters, some delay has been experienced. The total delay due to pilot measurement, filtering, and handover signaling may be on the order of several seconds. If the mobile station runs with a 50 km/h speed, it moves almost 14m during 1 second. If the size of a microcell is 100m, only a 7 second delay on handover is needed to go to the center of the cell.

8.4 MICRO- AND MACROCELLS AT DIFFERENT FREQUENCIES

A system with different cell layers at different frequencies is easier to manage since cell layers do not interfere with each other as much as when they are at the same frequency. The main interference comes from the adjacent channel spill-over. A drawback of this approach is large spectrum requirements since each cell layer requires its own frequency. For example, as was discussed in Chapter 6 (Figure 6.2), the WCDMA scheme implements three layers within 15 MHz.

In this section, first ACI attenuation masks for the system level simulations are generated by using the characteristics of a measured power amplifier IC. The capacity for the uplink and downlink of a HCS network is then evaluated. Impact of power control, handover, and channel spacing on the system capacity are also discussed.

8.4.1 Adjacent Channel Interference and Link Level Performance

Adjacent channel interference and link level performance degradation determine the impact of nonlinear power amplifiers on overall spectrum efficiency. As shown below, the increased adjacent channel interference will dominate the degradation of spectrum efficiency due to nonlinear power amplifiers.

In order to simulate a HCS network with the system level simulator, adjacent channel interference has to be modeled. Here, adjacent channel interference is modeled so that the output power of an interfering user is reduced by *adjacent channel attenuation.* Adjacent channel attenuation describes how much the power of the interfered user is attenuated if it is received at the adjacent channel. Adjacent channel interference attenuations are simulated with OQPSK chip modulation, bandlimited with square-root raised cosine pulse shaping and using the described power amplifier model. Roll-off factor is selected to be 0.20 in the simulations. Because the system includes different chip rates for macro- and microcells (2.5 and 5.1 Mchip/s), two separate ACI masks need to be created for system level simulations. These masks are created by backing off the power amplifier by 3.5 dB and by measuring the spectrum spreading power with the receiver square-root raised cosine filter also having a roll-off factor of 0.2. By sliding the receiver filter in frequency domain, continuous ACI masks are created for system level simulations, as shown in Figure 8.5. The backoff used corresponds to about 31% power amplifier efficiency [4]. Because the measured amplifier effects are quite equal to OQPSK and QPSK modulation methods [4], the same ACI masks can be used for the downlink system level performance simulations.

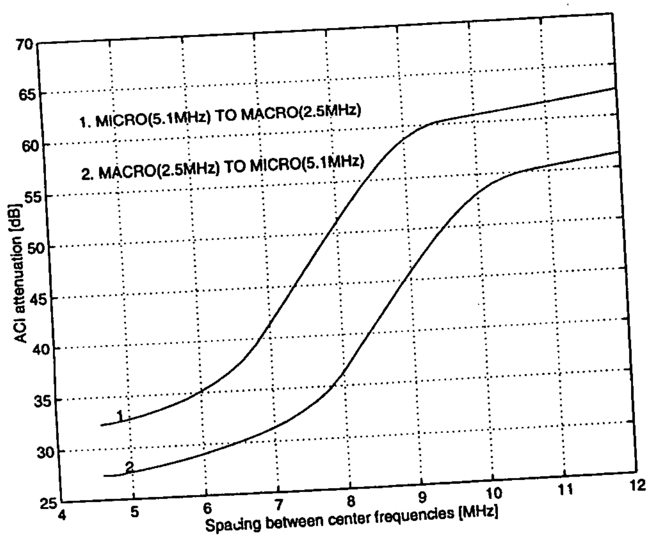


Figure 8.5 ACI masks for system level simulations.

Link level performance was also simulated with the same power amplifier model and with OQPSK and QPSK chip modulations, bandlimited with square-root raised cosine filter having a roll-off factor of 0.2. CDMA was found to be very robust against the nonlinearity of the power amplifier. Less than 0.2 dB link level performance losses occurred (both uplink and downlink) when the same output backoff (3.5 dB) as in the ACI mask definitions was used. This means that in a CDMA system, the spectrum mask requirements will limit the use of nonlinear amplifiers rather than link level performance losses. On the other hand, this means that if moderate ACI attenuation requirements can be tolerated in CDMA, the use of highly efficient power amplifiers becomes possible.

8.4.2 System Level Simulations of a HCS Network

In Figure 8.6, a HCS case with a hotspot of traffic in the middle of the area is shown. The high density of users located at the hotspot is served by microcells. The whole system area is covered by macrocells, which are used to provide continuous coverage.

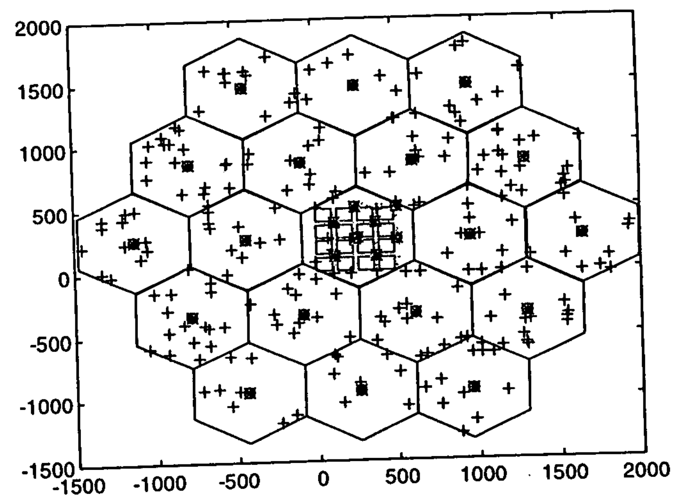


Figure 8.6 HCS environment where the macrocell base stations are on a hexagonal grid and the mobile stations (plus signs) are generated only to the streets. Eight microcell base stations are placed in the middle of the map (hotspot).

Figure 8.7 shows continuous coverage of microcell base stations. Here, both micro- and macrocell base stations provide continuous coverage. Fast moving mobile stations are handled by macrocell base stations, while high capacity terminals are served by microcell base stations.

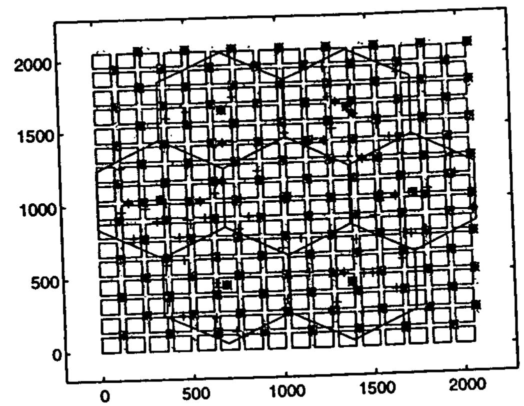


Figure 8.7 HCS environment with the full coverage of microcells.

System level simulations are performed at two HCS environments so that the hierarchy layers are separated in frequency. In the first case, the simulated system is covered by full coverage of macro- and microcell base stations. In this case, the number of macro- and microcell base stations are 7 and 128, respectively. In the second case, microcells are placed in a hotspot, while macrocells continuously cover the simulated area. The number of microcells is 8, and the number of macrocells is 19. In the simulation, macrocell base stations are placed in a hexagonal grid and the microcell base stations in a Manhattan-like grid. Microcell base stations are located in every second street intersection. All mobile stations are located in street canyons.

Instantaneous SIR is calculated by dividing the received signal by the interference and multiplying by the processing gain. In the uplink, it is assumed that a signal propagates from a mobile station to a base station via four equal strength Rayleigh faded paths. The observed base station adds together paths received from the observed user. It is supposed here that signals are combined coherently with maximal ratio combining. Maximal ratio combining is modeled by taking the sum of SNR values of each received path [5]. When simulating medium bit rate services, interference can be assumed to be Gaussian, since it is composed of transmissions to several independent users: signals from other users are just added together and seen as interference. If an interfering user was connected to the observed base station, the multiuser detection process is able to cancel the signal partially. The cancelled signal is multiplied by *(1–β)*, where β is *multiuser detection efficiency* and given as a parameter. SIR can be written as

$$SIR_{UL} = \frac{G_p \cdot S}{(1-\beta)\cdot I_{intra} + I_{inter} + N_0} \qquad (8.1)$$

Since the system is interference limited, thermal noise N_0 is assumed to be small and neglected. External interference I_{inter} consists of intercell and inter-hierarchy level interference. Inter-hierarchy level interference is multiplied by *adjacent channel attenuation* δ. Thus, I_{inter} is

$$I_{inter} = I_{intercell} + \delta \cdot I_{interlayer} \qquad (8.2)$$

In the downlink, signals from each base station to the mobile station propagate via four independent equal strength Rayleigh faded paths. Since macro diversity was considered at link level, the mobile station is able to receive all the paths from the base stations transmitting to it. Interference that propagates via the same paths than the desired signal is multiplied by an *orthogonality factor* α. This orthogonality factor states the orthogonality loss due to multipath propagation and energy losses due to limited receiver capacity. The orthogonality factor α is a simulation parameter since it varies according to radio environment, depending on the multipath delay spread (see Section 7.5.3). The transmission of the base stations that do not direct their transmission to the observed user contributes to the noise. SIR will then be

$$SIR_{DL} = \frac{G_p \cdot S}{(1-\alpha)\cdot I_{intra} + I_{inter}} \qquad (8.3)$$

In (8.3) I_{inter} consists of intercell and inter-hierarchy level interference, where interference from other hierarchy levels is multiplied by δ.

A performance measure for system simulations is outage percentage. A maximum of 5% of the obtained SIR values are allowed to be lower than the threshold SIR. SIR distributions are generated for both hierarchy layers and outage requirements that have to be fulfilled for both. Thus, the total outage requirement becomes even tighter than 5%. When the developed outage for one layer is 5%, the corresponding outage with the selected load will probably be less than 5% for the other layer. Simulation parameters are shown in Table 8.1.

8.4.3 Spectrum Efficiency Results

Spectrum efficiency results are shown in Table 8.2. Single layer micro- and macrocell capacities were simulated and used as references for HCS simulations. As seen previously, the capacity of the uplink with multiuser detection and antenna diversity is higher than the near orthogonal downlink capacity. In HCS simulations with full coverage of microcells, the microcell layer was loaded to 80% of a single layer microcell capacity. In case of a hot spot, macrocells were loaded to 80% of a single layer capacity. As seen from the results, co-existing layer capacity is almost the same as the single layer capacity, even when the interfering layer was loaded to 80% of its capacity. Even if the difference between pathloss of a microcell user and macrocell user may be several dozen decibels, power control with high dynamics makes the situation easy.

Capacities shown in Table 8.2 are obtained with high dynamic power control range. If the downlink power control range is set to 10 dB, a high difference between pathloss values cannot be compensated and outage becomes poor with the selected

loading. Outage requirements can be fulfilled if the load is decreased or if channel spacing is increased. Very high power control dynamics may be difficult to implement in the downlink. Therefore, power control dynamics and channel spacing should be selected so that performance versus bandwidth usage is optimized and implementation requirements fulfilled. As an example, power control dynamics could be 10 dB and channel spacing 1.6 MHz higher than in the reference case. Then, on average, a 15 dB higher difference on pathloss values can be compensated. Another way to increase the ACI attenuation is to raise the output backoff of the power amplifier. Unfortunately, this decreases the achievable power amplifier efficiency quite fast, as shown in Figure 8.3. If the number of microcells is low, then the intercell interference ($I_{intercell}$) becomes low and the HCS microcell capacity can be even higher than the reference capacity. This is the case if microcells are in a hotspot. The same effect occurs for macrocell, as can be seen in the case where microcells cover a large area and the number of macrocell base stations is low (seven base stations).

Simulation results shown in this chapter are different from the results shown in Chapters 5 and 7, as the selected parameter set and channel models were different.

Table 8.1
System level simulation parameters.

Parameter	Uplink	Downlink	Unit
Chip rate: - macrocell	2.5	2.5	MHz
- microcell	5.1	5.1	MHz
Channel separation	4.6	4.6	MHz
Handover margin	3	3	dB
Active set size	3	3	-
MUD efficiency: - 144 Kbps, micro	60	-	%
- 12 Kbps, macro	40	-	%
Orthogonality, - macro	-	68	%
- micro	-	62	%
ACI attenuation (δ): - Macro-to-micro	28	28	dB
- Micro-to-macro	32	32	dB
Power control step size	1	1	dB
Power control dynamics	80	10 and 80	dB
E_b/N_o, - 144 Kbps, micro	0.4	6.4	dB
- 12 Kbps, macro	5.7	6.4	dB

Table 8.2
Simulation Results [Kbps/MHz/cell]

	Uplink		Downlink	
	Micro	Macro	Micro	Macro
Hot spot	1138	204	317	175
Full	634	271	216	276
Reference	792	236	274	218

In [6], the case where adjacent channels overlap (totally or partially) was neglected. From the simulations, it is expected that a CDMA-based network could operate so that both hierarchy levels use the same frequency. This, however, requires very fast power control and handover, and a very wide power control dynamic range for both transmission directions. Since the simulations were performed as snapshot simulations, the mobile station did not move during the simulation, and effects due to mobility were neglected. Mobility may destroy systems where the two hierarchy levels operate with the same frequency. Consider a case where a mobile station connected to the macrocell moves rapidly to the area of a microcell. The handover process is so slow that a mobile station runs deep inside microcell area before it is handed to the microcell base station. The only solution is very fast power control. If the mobile stations connected to the microcell base station are able to adjust their uplink powers fast enough, the system will not crash. The mobile stations connected to macrocells should have such a fast downlink power control that they can adjust downlink power in a way that the connection does not drop due to increased interference power from the microcell. In addition, the power control range for downlink should be so large that mobile stations can compensate for increased interference. The conclusion from system simulations is that hierarchical cell structure networks are feasible with CDMA, but some implementation problems may occur.

REFERENCES

[1] Andermo, P. G., "System Flexibility and its Requirements on Third Generation Mobile Systems," *Proceedings of PIMRC'92*, Boston, Massachusetts, USA, 1992, pp. 1–3.

[2] Shapira, J., and R. Padovani, "Spatial Topology and Dynamics in CDMA Cellular Networks," *Proceedings of VTC'92*, Denver, Colorado, USA, 1992, pp. 213–216.

[3] Shapira, J., "Microcell Engineering in CDMA Cellular Networks," *IEEE Transactions on Vehicular Technology*, Vol. 43, No. 4, November 1994, pp. 817–825.

[4] Lilja, H., "Characterizing the Effect of Nonlinear Amplifier and Pulse Shaping on the Adjacent Channel Interference with Different Data Modulations," *Licentiate Thesis*, Oulu University, Finland, 1996.

[5] Proakis, J. G., Digital Communications, New York: McGraw-Hill, 1995, pp. 778–785.

[6] Hämäläinen, S., H. Lilja and, J. Lokio, "Performance of a CDMA Based Hierarchical Cell Structure Network," Proceedings of PIMRC'97, Helsinki, Finland, September 1997, pp. 863–866.

Chapter 9

TIME DIVISION DUPLEX DS-CDMA

9.1 INTRODUCTION

In time division duplex (TDD), the uplink and downlink transmissions are time multiplexed into the same carrier, in contrast to frequency division duplex (FDD), where uplink and downlink transmissions occur in frequency bands separated by the duplex frequency. Figure 9.1 illustrates the principles of TDD and FDD.

Section 9.2 introduces second generation TDD systems: DECT, PHS, and CT2. Section 9.3 presents the motivation for using TDD for third generation systems.

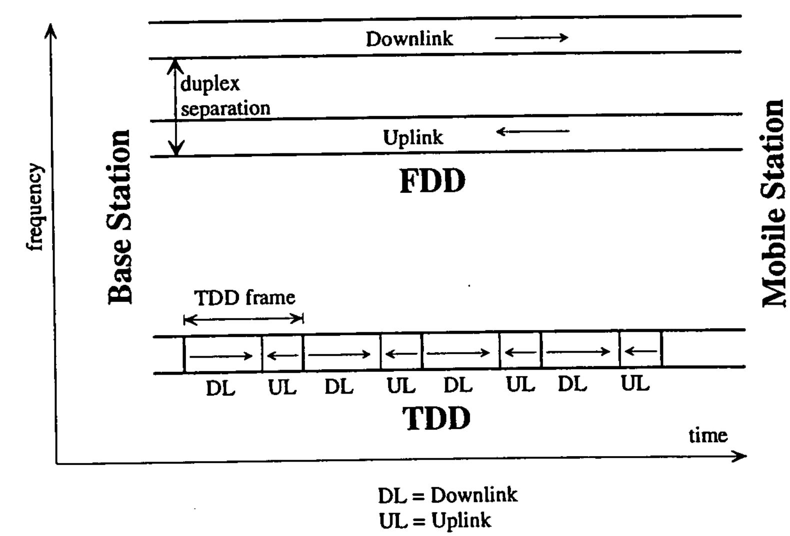


Figure 9.1 TDD and FDD principle.

In Section 9.4, interference aspects in TDD are considered. Intercell and intracell interference problems in asymmetric TDD-CDMA system are identified, as well as interference between operators. TDD-specific air interface design aspects, in particular the frame length and the receiver structure, are discussed in Section 9.5. A third generation TDD-CDMA system as an extension to an FDD-CDMA system is presented in Section 9.6. The performance of the system is analyzed for a WWW browsing traffic model. Section 9.7 summarizes other proposed wideband CDMA-TDD systems. TDMA-based third generation TDD proposals such as in [1] are described in Chapter 1.

9.2 SECOND GENERATION TDD SYSTEMS

Examples of second generation TDD systems are Digital European Cordless Telephone (DECT), Personal Handy Phone System (PHS), and CT2. The main parameters of these systems are listed in Table 9.1. These systems are intended for a low tier radio environment, mainly for indoor operation. A common feature of second generation TDD systems is no or very little channel coding, which restricts the performance to low mobility radio environments. TDD systems have not gained as much market support as the second generation FDD technologies (GSM, IS-95, PDC, and US-TDMA). The main reason for this seems to be the limited mobility and coverage provided by the TDD systems.

The DECT standard was adopted by ETSI in 1992, and during 1993 the first generation of DECT products were commercially launched. In addition to speech services, DECT provides data services with bit rates up to 512 Kbps half duplex and 256 Kbps full duplex. DECT also provides interworking with GSM.

PHS was developed in Japan in the early 1990s. Commercial operation was launched in 1995. An interesting feature specific to PHS is that direct mobile-to-mobile calls are allowed using the lower end of the frequency band (1895 to 1898 MHz). In the United States, PHS was merged with the WACS system in 1994. This formed Personal Access Communications System (PACS), which has also an FDD mode.

Table 9.1
TDD System Parameters

	DECT	PHS	CT2
Multiple access	FDMA/TDMA	FDMA/TDMA	FDMA
Frequency band	1880 – 1900 MHz	1895 – 1918 MHz	864 – 868 MHz
Carrier bandwidth	1.728 MHz	300 kHz	100 kHz
Carrier bit rate	1.152 Mbps	384 Kbps	72 Kbps
Frame length	10 ms	5 ms	2 ms
Timeslots/frame	24	8	2
Modulation	GMSK	DQPSK	GFSK
Speech coding	ADPCM 32 Kbps	ADPCM 32 Kbps	ADPCM 32 Kbps
Max. transmitter power	240 mW	10 mW	10 mW
Power control	No	Yes	No

9.3 REASONS TO USE TDD

This section discusses the reasons for selecting TDD as a duplex scheme for a third generation CDMA system.

9.3.1 Spectrum Allocation

Providing high bit rate services, even with high spectral efficiency, requires a large bandwidth. Additionally, the services required beyond the year 2000 are not yet clearly defined. Efficient and flexible utilization of all the available bandwidth, including the TDD band, is essential for viable third generation mobile radio systems. Figure 9.2 and Table 9.2 show spectrum for IMT-2000 according to WARC'92. Since the IMT-2000 spectrum allocation is asymmetric, it supports both FDD- and TDD- based systems. In Europe, ERC has designated these frequencies for UMTS [2]. Even though not specifically decided, frequency bands 1900 – 1920 and 2010 – 2025 MHz are unpaired bands and could be used for TDD applications. Japan has adopted a similar frequency plan for IMT-2000.

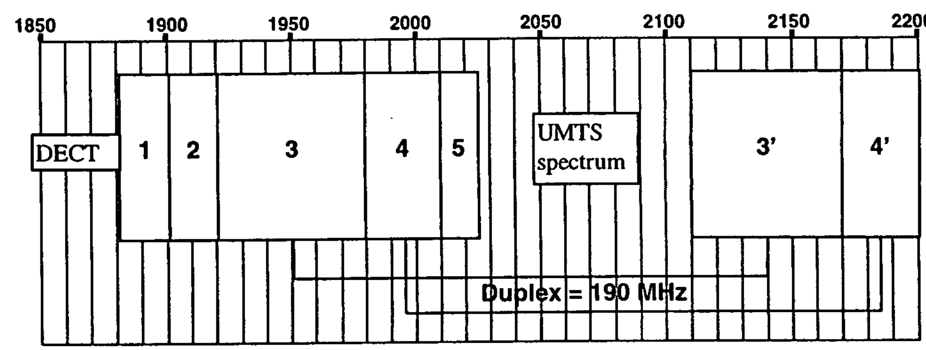


Figure 9.2 Proposed spectrum allocation for UMTS. Frequency band 1 is currently used by DECT.

Table 9.2
Proposed Spectrum Allocation for UMTS

Band	Frequency (MHz)	Bandwidth (MHz)	Allocation
2	1900 – 1920	20	UMTS terrestrial applications (TDD)
3	1920 – 1980	60	UMTS terrestrial applications (FDD)
4	1980 – 2010	30	UMTS satellite component (FDD)
5	2010 – 2025	15	UMTS terrestrial applications (TDD)
3'	2110 – 2170	60	UMTS terrestrial applications (FDD)
4'	2170 – 2200	30	UMTS satellite component (FDD)

9.3.2 Asymmetric Services

Services such as the Internet, multimedia applications, and file transfers often set different capacity requirements for the uplink and downlink. The utilization of a TDD frequency band is not fixed between the uplink and downlink (unlike with FDD) and this flexibility in resource allocation can be used if the air interface design is flexible enough. This is one motivation for considering the TDD extension for FDD-CDMA–based full coverage third generation mobile radio systems.

9.3.3 Reciprocal Channel in Uplink and Downlink

In FDD operation, the uplink and downlink transmissions are separated by a duplex separation. Since fast fading due to multipath propagation depends on the frequency, it is uncorrelated between the uplink and downlink. The FDD transmitter cannot predict the fast fading that will affect its transmission.

In TDD operation, the same frequency is used for both the uplink and the downlink. Based on the received signal, the TDD transmitter is able to know the fast fading of the multipath channel. This assumes that the TDD frame length is shorter than the coherence time of the channel. This assumption holds if TDD mobiles are slowly moving terminals. The reciprocal channel can then be utilized for:

- Open loop power control;
- Spatio-temporal transmission diversity (adaptive antennas for transmission [3], pre-RAKE [4–6]).

With open loop power control, the need for power control signaling is reduced compared to closed loop power control. Closed loop power control signaling also introduces some delay and is subject to errors, which is not the case with open loop power control. In order to have fast enough open loop power control, the TDD frame must be short enough. According to [7], Doppler frequencies up to 80 Hz (43 km/h at 2-GHz carrier frequency) can be supported with a very small degradation if the uplink part of the TDD frame length is 1.5 ms. If the TDD system is intended only for slowly moving terminals, then longer TDD frames could also be used. With open loop power control, the interference situation at the receiver is not known by the transmitter, only the signal level is known. Changes in the interference level must be signaled.

Transmission diversity can be utilized with diversity antennas (space domain diversity) or with pre-RAKE (time domain diversity). In selection diversity combining, the receiver measures the received signal from diversity antennas and selects the best antenna for reception. Antenna diversity techniques are easily applied at the base stations but those receiver techniques are not suited for small handheld terminals. In order to achieve antenna diversity in the downlink, transmission diversity is utilized at the base station. Based on uplink reception, the best antenna can be selected for downlink transmission in TDD.

In FDD-CDMA transmission, a RAKE receiver is used to collect the multipath components and to obtain multipath diversity. The optimal RAKE receiver is a matched

filter to the multipath channel. If the transmitter knew the multipath channel, it could apply RAKE in the transmitter (pre-RAKE). Transmission would be such that the multipath channel would act as a matched filter to the transmitted signal. In the receiver no RAKE would be needed (i.e., multipath diversity could be obtained with a one-finger receiver). However, it should be noted that only a little multipath diversity may be available in indoor propagation environments, as shown in Chapter 7 with ITU channel models. Indoor and microcell environments are the most probable application areas for TDD communication. Therefore, antenna diversity transmission will be a more attractive diversity technique for TDD operation than the multipath diversity technique.

If such a TDD proposal, which has different solutions for the uplink and downlink (where uplink uses a single wideband carrier and downlink a multicarrier approach) is applied, channel reciprocity cannot be utilized. The effect of such a structure depends on the environment, but, as calculated in Chapter 7, the coherence bandwidth cannot be guaranteed to be so high that similar fading characteristics could be obtained for both uplink and downlink.

9.4 PROBLEMS WITH TDD-CDMA

This section introduces the problems encountered by TDD-CDMA systems and presents the disadvantages of using such a duplex method.

9.4.1 Interference From TDD Power Pulsing

If fast power control frequency with open loop is desired to support higher mobile speeds, then short TDD frames must be used. The short transmission time in each direction results in the problems listed below:

- Audible interference from pulsed transmission both internally in the terminal and to the other equipment. Generated pulsing frequency in the middle of voice band will cause problems to small size speech terminal design where audio and transmission circuits are relatively close to each other and achieving the needed isolation is costly and requires design considerations. At high power levels this may not be achievable at all.
- Base station synchronization requirements are tight and more overhead must be allocated for guard times and also for power ramps as EMC requirements limit the ramping speed.
- Fast ramping times set tighter requirements to the components (e.g., to the power amplifier).

Lower pulsing frequency, say, 100 Hz (i.e., a TDD frame of 10 ms), results in less audible pulsing but limits the maximum tolerable mobile speeds. In the TDD-CDMA in [8,9], the uplink slot and the downlink slot are both 0.625 ms, resulting in an audible interference at 800 Hz.

9.4.2 Intracell and Intercell Interference Between Uplink and Downlink

In CDMA systems, the SIR may be quite low (e.g., below −15 dB) at carrier bandwidth. After despreading, the SIR is improved by the processing gain. In TDD systems, a transmitter located close to a receiver may block the front end of the receiver, since no RF filter can be used to separate uplink and downlink transmission as in FDD operation. This blocking may happen even if the transmitter and receiver are not operating in the same frequency channel but if they are operating in the same TDD band. In that case, the processing gain at baseband does not help since the signal is already blocked before baseband processing. These interference problems with TDD operation are considered in this section.

Within one TDD-CDMA cell, all users must be synchronized and have the same time division between uplink and downlink in order to avoid interference between uplink and downlink. This time division is based on the average uplink and downlink capacity need in that particular cell. Each user then applies multirate techniques to adapt its uplink and downlink capacity needs to the average need in that cell. The same time division must be applied to all carriers within one base station. If the base station transmits and receives at the same time as adjacent carriers, it would block its own reception.

Asymmetric usage of TDD slots will impact the radio resource in neighboring cells. This scenario is depicted in Figure 9.3 and the resulting signal to adjacent channel interference ratio is calculated in Table 9.3. Intercell interference problems occur in asymmetric TDD-CDMA if the asymmetry is different in adjacent cells even if the base stations are synchronized. MS2 is transmitting at full power at the cell border. Since MS1 has different asymmetric slot allocation than MS2, its downlink slots received at the sensitivity limit are interfered with by MS1, causing blocking. On the other hand, since BS1 can have much higher effective isotropically radiated power (EIRP) than MS2, it will interfere with BS2 receiving MS2. It is difficult to adjust the asymmetry of an individual cell in a network due to interference between adjacent cells. If TDD-CDMA cells are located adjacent to each other, offering a continuous coverage, then synchronization and asymmetry coordination between these cells is required. This ensures that the near-far problems of interference between mobiles in adjacent cells can be controlled. Another scenario, where the previously described blocking effect clearly exists, is if TDD operation were also allowed in the FDD band.

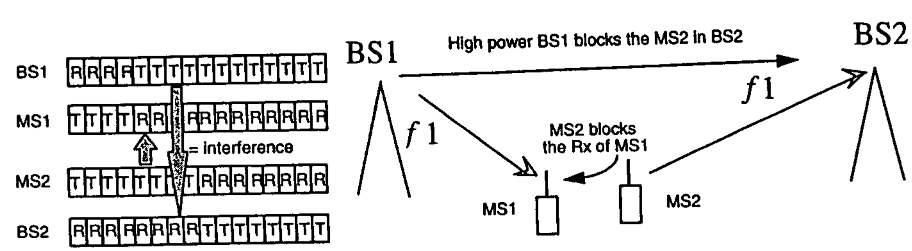


Figure 9.3 TDD interference scenario.

Table 9.3
Adjacent Channel Interference Calculation

BTS transmission power for MS2 in downlink 1W	30 dBm
Received power for MS1	−100 dBm
Adjacent channel attenuation due to irreducible noise floor	50 to 70 dB
Signal to adjacent channel interference ratio	−60 to −80 dB

9.4.3 Inter-Operator Interference with Continuous Coverage

If there are several operators offering the service in the same geographical area in the TDD band, base station synchronization for the different operators is required and asymmetry flexibility between uplink and downlink becomes considerably more difficult. Asymmetric TDD-CDMA systems are therefore not well suited if several operators share the spectrum and the same area. In those cases, TDD-TDMA systems are better off since the signal is concentrated in the time domain. In TDD-TDMA, interference can be averaged with time hopping or avoided with dynamic channel allocation techniques not applicable in CDMA systems. A difficult problem for operation in an unlicensed band is the case where different operators use different multiple access techniques in the same TDD band.

Adjacent channel interference between operators may also cause problems. As the terminals have limited dynamic range and neighboring channel filtering capability, adjacent channel interference in this kind of uncoordinated operation may prove to be very severe. The power differences between the transmission of the desired base station and the interfering mobile transmitting at the same time as adjacent carriers can block the receiver terminal's A/D converters and can also cause problems in the RF components.

Assuming the 15-MHz spectrum for uncoordinated TDD, a chip rate of 4.096 Mcps can then accommodate only three carriers. If symmetric 2.0 Mbps should be supported with a TDD network, a double chip rate of 8.192 Mcps would be required.

9.4.4 Synchronization of Base Stations

Intercell and inter-operator interference problems are present in a TDD-CDMA system if the base stations are not synchronized. Synchronization is therefore desirable for TDD-CDMA. Synchronization accuracy must be at the symbol level but not at the chip level. It can be achieved with, for example, GPS receivers at the base station or by distributing a common clock with extra cabling. These methods increase the cost of the infrastructure.

9.5 TDD-CDMA AIR INTERFACE DESIGN

The specific aspects of a TDD-CDMA system air interface design are presented in this section, with particular attention to frame design.

9.5.1 TDD Frame Length

Since within coherence time the channel for uplink and downlink is the same, there is no need for closed loop power control. However, the frame structure limits the maximum command rate for the power control. The mobile station measures the received power during the downlink transmission and determines the uplink transmission power, which is applied during the uplink transmission. Thus, if the frame length is 10 ms, the command rate is 100 Hz. For power control, no feedback can be provided during the uplink part: thus, it is desirable that the channel is constant during the uplink part. If symmetric division between uplink and downlink is assumed, the uplink part is 5 ms in this example. Coherence time depends on the Doppler frequency and is shown in Table 9.4 for 2-GHz carrier frequency. Coherence time should be clearly longer than the uplink part of the TDD frame for power control to perform adequately. With a 10 ms TDD frame (i.e., 5 ms uplink part in symmetric allocation), open loop power control is effective for mobile speeds of up to about 10 to 20 km/h.

Table 9.4
Coherence Times for Different Mobile Speed

Mobile station speed (km/h)	Doppler frequency at 2 GHz (Hz)	Coherence time (ms)
5	9	108
10	18	54
20	37	27
30	56	18
40	74	14
50	93	11
80	148	6.8
100	185	5.4

9.5.2 Hardware Requirements

Since the channel is reciprocal, it is possible to simplify the receiver while still maintaining effective diversity techniques against fading. In TDD-CDMA, pre-RAKE could be applied either to downlink or to uplink. If pre-RAKE is utilized in uplink, simple TDD base stations could be built. This option is attractive if all the terminals support both FDD and TDD operation, and they must therefore have a RAKE receiver for FDD operation. In that case, TDD base stations could be made very simple when no RAKE is needed. Also, multiuser detection algorithms at the base station would be simpler if there is only one multipath component per user to be tracked. If TDD-only terminals are used, then pre-RAKE could be applied to the base station to reduce the complexity of TDD terminals. Pre-RAKE cannot be applied to both transmission directions at the same time.

Downlink performance could be improved with transmission antenna diversity at the base station. The downlink transmission antenna is determined based on the uplink reception. Also, downlink beamforming with adaptive antennas is easier to utilize in TDD than in FDD, but in typical TDD environments (like indoor), users may be difficult to separate by the direction of arrival.

In a dual-mode FDD/TDD terminal, additional complexity is concentrated in the RF filtering, compared to a FDD only terminal. Modifications at the baseband parts are minor compared to changes in the RF parts.

9.6 TDD-CDMA EXTENSION TO FDD-CDMA SYSTEM

A TDD-CDMA extension to a FDD-CDMA based third generation cellular system, as presented in [10], is considered in this section. A stand-alone TDD-CDMA system providing full coverage is not considered because of possible interference problems in TDD operation. A system scenario with both TDD and FDD is presented, where frequency bands allocated to TDD operation can be effectively utilized in complementing the services and increasing the capacity provided by a wide area coverage FDD system.

It is assumed that the FDD-CDMA system is used to provide both wide area speech and data services. The FDD-CDMA operator offers public access to the network for all IMT-2000 users. The TDD-CDMA with limited coverage is used as an extension to the FDD-CDMA. The TDD system typically covers hotspots with high capacity requirements, such as office buildings, airports, and hotels. These TDD systems are operated either by the same public operator as the FDD system or by private operators. such as companies in their own premises or service providers providing the access in office buildings shared by several companies. The TDD system is not intended to offer continuous coverage, and systems in different locations can operate independently of each other. The system scenario is shown in Figure 9.4.

9.6.1 TDD-CDMA System Description

The TDD-CDMA extension is based on the FDD-CDMA system. When the parameters of the TDD extension (bandwidth, frame length, chip rate) are aligned with the FDD, the implementation of FDD-TDD terminal is easier.

It is assumed here that CDMA terminals support both FDD and TDD operation. By employing pre-RAKE in mobile stations, the complexity requirements of TDD-CDMA base stations are reduced (note that mobile stations already have RAKE receiver due to the FDD operation). The simplified architecture for the TDD-CDMA mobile terminal is shown in Figure 9.5. During the downlink (three switches shown in Figure 9.5 are up), the received signal is despread and this correlation process is followed by a RAKE filter. Pilot symbols within the TDD frame are used to form a channel estimate. During the uplink (switches down), the channel estimate obtained from the downlink is used in the pre-RAKE filter, and the channel effectively forms a matched filter to the transmitted signal.

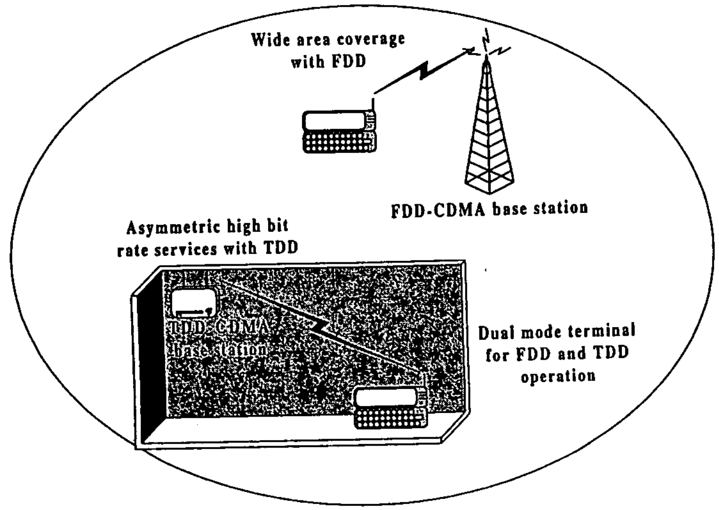


Figure 9.4 TDD-CDMA extension to full coverage FDD-CDMA system.

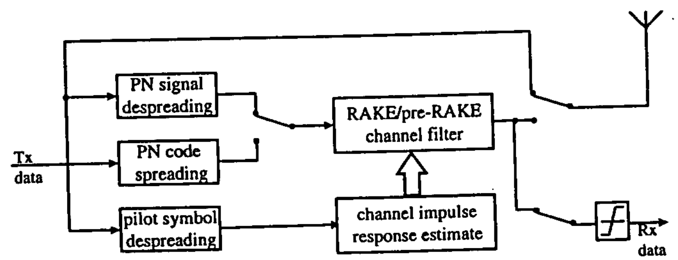


Figure 9.5 Simplified block diagram of TDD-CDMA mobile terminal architecture.

The capacity of the TDD system is divided asymmetrically between the uplink and downlink to support multimedia services. This asymmetry can be adjusted to match the uplink and downlink requirements for the TDD-CDMA cell. All downlink channels will be synchronized to obtain good orthogonality between users. The TDD-CDMA asymmetric frame structure is shown in Figure 9.6. The simulation results in this chapter are obtained with a frame length of 20.5 ms, and it comprises an uplink burst, two guard time bits, and a downlink burst. The guard times in a TDD system must accommodate the maximum round trip delay time in the cell as well as the hardware delays.

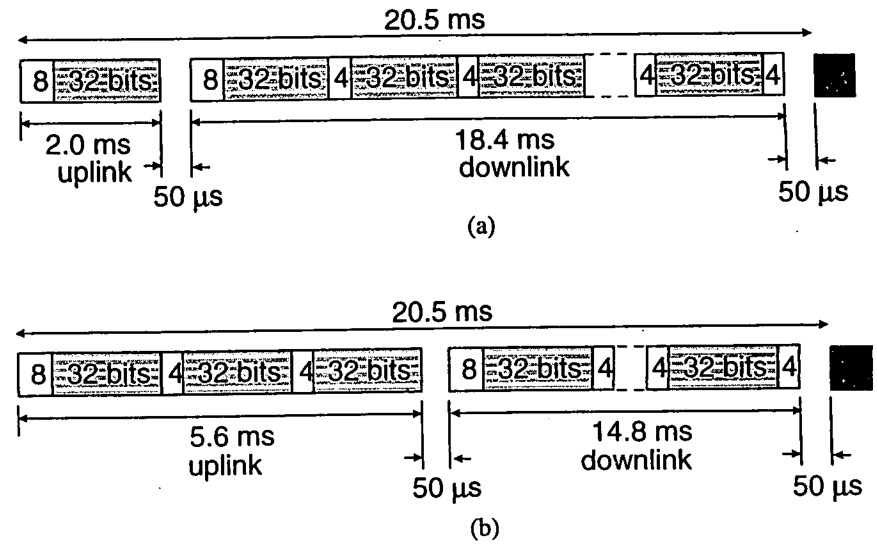


Figure 9.6 TDD-CDMA asymmetric frame in pre-RAKE simulation: (a) uplink/downlink load = 1/10 and (b) uplink/downlink load = 3/8.

A frame which has a downlink capacity of 10 times that of the uplink is shown in Figure 9.6(a) and corresponds to the WWW browsing model described in Table 9.5. The TDD-CDMA air interface is flexible and will allow extra timeslots to be added to the uplink at the expense of downlink timeslots, as shown by Figure 9.6(b).

Different TDD asymmetries between uplink and downlink in near or adjacent cells can lead to high intercell interference levels. To combat this, such cells must either have the same asymmetry and be synchronized, or be synchronized with certain timeslots being replaced by guard times (which is effectively a loss of capacity). This is based on the assumption that the cell reuse efficiency would be 100%.

The beginning of both the uplink and downlink for each frame contains an 8-bit header. Four of these bits are pilot bits, and four are for control. Higher level control can be provided by arranging the frames into superframes. The frame is divided into eleven 32 bit timeslots, each being preceded by four pilot bits. On the synchronous downlink, the pilot bits are transmitted at 100% of the transmitter power and provide training for RAKE (and pre-RAKE); they also aid the synchronization process. Pilot overhead is equivalent to 12.5% of total capacity. One of the downlink control bits can be used for closed loop power control. Only a single bit per frame is necessary to provide fine adjustments. The fast open loop power control is achieved due to the correlated nature of the uplink and downlink impulse responses, but the maximum

power control frequency is limited to 50 Hz by the TDD frame rate. Thus, fast moving terminals are not supported by the TDD-CDMA system since they will increase the near-far effect.

9.6.2 WWW Browsing Session Performance with TDD Extension

The simulation parameters for the TDD-CDMA extension are given in Table 9.5. Simulations have assumed perfect chip synchronization. The frame structure shown in Figure 9.6(a). is used in the simulations. The WWW model gives rise to a nonuniform document inter-arrival rate, and so, transmission buffers must be used. We assume that the document arrival at the transmitter takes zero time, but that its transmission is limited by the data rate provided by the communication link.

Table 9.5
TDD-CDMA System Simulation Parameters

Max. symbol rate	20 Kbps
Nominal uplink bit rate	1.56 Kbps
Downlink bit rate/channel	15.61 Kbps
Chip rate	5.12 Mcps
Code length	256
Modulation	BPSK
Channel coding	None
Doppler rate	10 Hz classic
Multipaths	4 path (equal powers)
Antenna configuration	Single antenna
RAKE/pre-RAKE	Implemented without pre-RAKE prediction
Multiuser detection	Not implemented
Traffic parameters	
Mean downlink WWW document arrival rate	60 seconds (Poisson distribution)
Mean document size	30 Kbyte (Rayleigh distribution)

Figure 9.7 shows theoretical results based on the downlink for an uncoded system with a perfect channel estimate applied to a four-tap RAKE filter. The number of users for this system is not the same as the number of channels utilized. The average number of channels required per user for the WWW model is about 1/4 so there is only one channel required for about every four users. The "user SNR" refers to the mean received SNR (pre-correlation) seen by each user, excluding any multiuser interference. The processing gain of code length of 256 is equal to 24.1 dB.

Figure 9.8 shows the simulated error rates for a different number of users. Here the RAKE filter uses the channel impulse response estimate obtained by averaging the four pilot bits provided for each of the downlink timeslots. No other filtering or signal processing was applied to this signal and the performance could be further improved. The last channel estimate provided by the downlink is used for the pre-RAKE filter since this will be highly correlated with the uplink impulse response at low Doppler spreads. No RAKE filter is implemented in the base station, and only a single antenna is considered in these simulations.

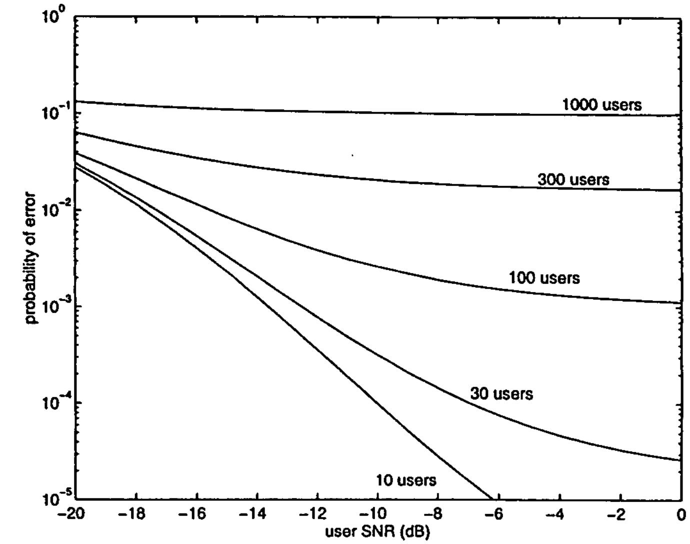


Figure 9.7 Error probability for downlink with perfect channel estimate and uniform channel loading. Processing gain is 24.1 dB.

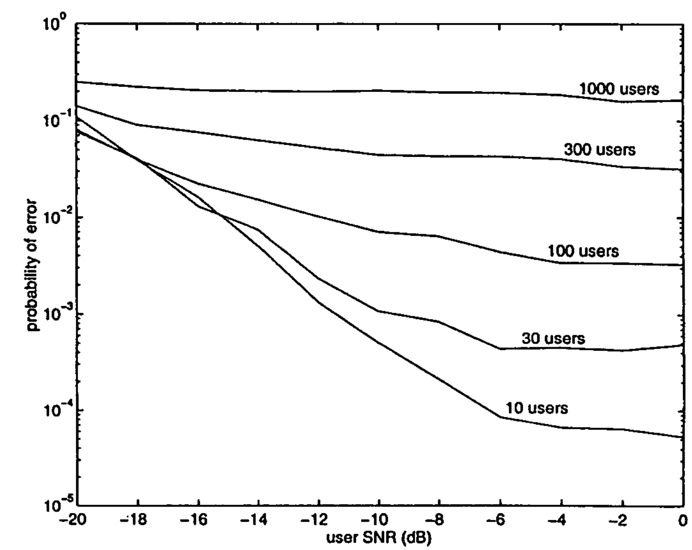


Figure 9.8 Simulated downlink error probability for the WWW service model using channel estimates derived from the pilot bits. Processing gain is 24.1 dB.

Clearly, the simulated results are poorer than the theoretical ones, which represent a lower bound. There are two reasons for degradation. First, theoretical results assume a constant load for each user, whereas the load in the simulated results is governed by the Poisson document inter-arrival process. This will cause some peaks in the load and corresponding peaks in error rate. Secondly, theoretical results assumed perfect channel estimation, whereas the simulated results must derive the channel estimate from the pilot symbols.

9.7 WIDEBAND TDD-CDMA SYSTEMS

In this section Cylink TDD-CDMA and third generation TDD-CDMA proposals are presented.

9.7.1 Cylink S-CDMA TDD for PCS

One of the first concepts to use TDD operation with CDMA is Cylink's S-CDMA [11,12] intended for the US ISM band at 902-928 MHz. The concept has the parameters listed in Table 9.6.

Table 9.6
Cylink S-CDMA TDD for PCS system concept parameters

Modulation	DS/QPSK
Voice coding	32 Kbps ADPCM
Data transmission	4.8 Kbps signaling channel
Processing gain	32
Synchronization	Network, both downlink and uplink
Users per carrier	Max 16
QPSK Symbol burst rate	1.536 Msymbol/s

Source: [13].

This proposal contains the basic features of a cellular system. Base stations transmit broadcast information as part of their TDD burst and thus allow the mobiles to acquire frequency and time synchronization and receive general base station information. The concept uses sequences optimized in terms of autocorrelation, and all users use the same sequence within the cell. For this kind of operation, in addition to power control, timing control by the base station towards the mobiles is also required to maintain the achievable (near) orthogonality between users. For both network and user, synchronization accuracy of the order of one chip interval (651 ns) is required. The operation environment is considered to be indoor and indoor-to-outdoor with velocity below vehicle speeds (pedestrian environment), where delay profiles remain relatively small and do not change very rapidly due to low mobile speeds.

Base stations use preamble (pilot) data in the beginning of the common broadcast part to allow synchronization and acquisition of the signal. The use of base station transmission diversity is also included.

9.7.2 Third Generation Wideband CDMA TDD Schemes

A 5-MHz TDD-CDMA cellular system with synchronized base stations has been proposed in [8,9]. This system has also been adopted in ARIB as part of the wideband CDMA proposal and has also been the basis for the WCDMA TDD mode in ETSI and cdma2000 TDD mode in TR45.5 in the United States. The ETSI TDD scheme has also been influenced by the TD-CDMA scheme described in Chapter 1. We expect further harmonization between the ARIB and ETSI wideband CDMA TDD modes in the future.

The system parameters are described in Table 9.7. Time division between uplink and downlink is proposed to be fixed (i.e., asymmetric capacity allocation is not supported). Diversity scheme in base stations is used both in uplink and downlink by utilizing the reciprocal channel. Because TDD uses only one radio frequency band and transmission diversity at the base station can be employed, space diversity is not that important at the mobile station and thus the mobile station can be made smaller. TDD-CDMA systems can also perform effective transmission power control with only the use of open loop control. A single antenna and no duplexer are the merits of TDD-CDMA for terminal size and battery life. This system is especially effective in an environment such as microcellular systems or indoor office systems, where the spreading bandwidth is narrower than the coherent bandwidth, and path diversity gain by itself is small.

Table 9.7
System Parameters of Wideband TDD-CDMA

Access method	DS-CDMA/TDD
Minimum frequency band	5 (10/ 20) MHz
Chip rate	4.096 (8.192/16.384) Mcps
Frame length	10 ms = 0.625 ms × 8 slots × 2 (TDD)
Slot length	0.625 ms
Transmission power control	Open loop control + slow compensation control by control channel
Diversity	Mobile station: RAKE Base station: transmission/reception space diversity + RAKE
Modulation and Demodulation	Data: QPSK/pilot symbol coherent detection Spreading: BPSK
Error correction	Inner: K = 7 R = 1/2 convolutional code + Viterbi soft decoding Outer: R = 4/5 Reed-Solomon code
Inter-BS synchronization	Synchronization
Maximum user information transmission rate	144 Kbps (5 MHz bandwidth, vehicular) 2 Mbps (20 MHz bandwidth, indoor office)

For wideband TDD-CDMA, experimental equipment has been developed in order to evaluate such fundamental characteristics as reception diversity at base station, transmission diversity at base station, and open loop power control. The specifications of the TDD-CDMA testbed are shown in Table 9.8. Spreading bandwidth is 5 MHz for chip rate of 4.096 Mcps. The maximum number of space diversity branches is four at base station, and the maximum number of RAKE fingers is six. This testbed consists of one base station and two mobile stations (8-Kbps transmission user and 144-Kbps

transmission user). Simultaneous communication of 8 Kbps transmission user and 144 Kbps transmission user is possible.

Table 9.8
Main Testbed Specifications

Chip rate	4.096 Mcps (5-MHz bandwidth)
Power control	Open loop control + slow compensation control by control channel Closed loop control
Diversity	Mobile station: 1 branch antenna and 6 fingers RAKE combining Base station: 1 to 4 branch antenna diversity and 6-finger RAKE combining
FEC	Inner: K = 7, R = 1/2 convolutional code Viterbi soft decoding 10/40/80/160/320 ms interleaving Outer: R = 9/10 Reed-Solomon code 40/80/160/320 ms interleaving

9.7.2.1 *IMT-2000 TDD Proposals*

Four wideband CDMA TDD schemes have been submitted to ITU RTT evaluation: ARIB WCDMA TDD, ETSI WCDMA TDD, cdma2000 TDD, and a time division synchronous code division multiple access (TD-SCDMA) proposal from China. The ARIB WCDMA TDD scheme is very similar to the above described wideband CDMA TDD scheme.

ETSI WCDMA TDD: In the ETSI wideband CDMA TDD mode, a wideband CDMA carrier with chip rate of 4.096 Mcps is divided between uplink and downlink in time. The frame length is as in the FDD mode, currently 10 ms, and the number of time slots per frame is 16. Figure 9.9 shows an example TDD frame structure with one switching point for uplink / downlink separation within a frame. Another option under study is to use multiple switching points. As shown in Figure 9.9, a burst consists of three parts (data block - midamble - data block). The TDD mode uses QPSK data modulation and currently employs fixed spreading factor. In addition, variable spreading factor is under study.

Cdma2000 TDD: The frame length of the cdma2000 TDD mode is 20 ms and it is divided into sixteen 1.25-ms slots. Every other slot is for TX and every other for RX. Similar to the FDD mode, the TDD mode has multicarrier and direct spread option.

TD-SCDMA: The TD-SCDMA concept has a chip rate of 1.1136 Mcps and a carrier spacing of 1.2 MHz. The frame length is 5 ms and each frame is divided in 8 time slots (4 for the uplink and 4 for the downlink). One time slot can accommodate 16 CDMA codes.

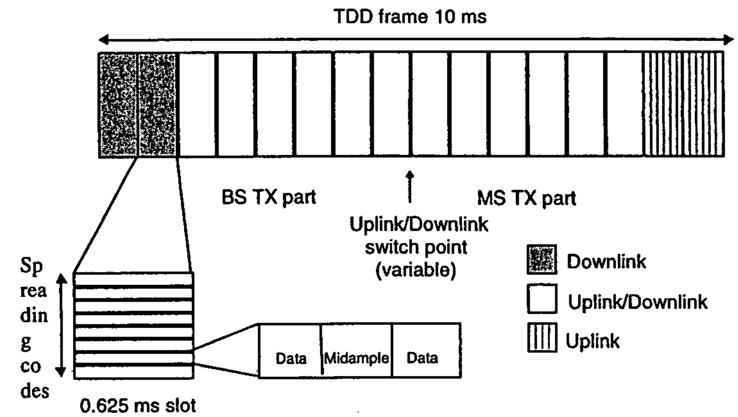


Figure 9.9 An example TDD frame structure.

REFERENCES

[1] Klein, A., R. Pirhonen, J. Sköld, and R. Suoranta, "FRAMES Multiple Access Mode 1 – Wideband TDMA with and without Spreading," *Proceedings of Personal, Indoor and Mobile Radio Communications (PIMRC'97)*, Vol. 1, September 1997, Helsinki, Finland, pp. 37–41.

[2] European Radiocommunications Committee ERC Decision of 30 June 1997 on the frequency bands for the introduction of the Universal Mobile Telecommunications System (UMTS), ERC/DEC/(97)07, can be found from ERO website http://www.ero.dk/.

[3] Litva, J., T. Lo, *Digital Beamforming in Wireless Communications*, Boston: Artech House, 1996.

[4] Povey, G. , "Frequency and Time Division Duplex Techniques for CDMA Cellular Radio," *Proceedings of 3rd International Symposium on Spread Spectrum Techniques & Applications (ISSSTA'94)*, Oulu, Finland, 1994, pp. 309–313.

[5] Povey, G., "Capacity of a Cellular Time Division Duplex CDMA System," *IEE Proc.-Commun.*, Vol. 141, No. 5, October 1994, pp. 351–356.

[6] Esmailzadeh, R., and M. Nakagawa, "Pre-RAKE Diversity Combination for Direct Sequence Spread Spectrum Mobile Communications Systems," *IEICE Transactions on Communications*, Vol. E76-B, No. 8, August 1993, pp. 1008–1015.

[7] Esmailzadeh, R., M. Nakagawa, and E. Sourour, "Time-Division Duplex CDMA Communications," *IEEE Personal Communications*, April 1997, pp. 51–56.

[8] Hayashi, M., K. Miya, O. Kato, and K. Homma, "CDMA/TDD Cellular Systems utilizing a Base-Station Based Diversity Scheme," *Proceedings of 45th Vehicular*

Technology Conference (VTC'95), July 1995, Chicago, Illinois, USA, Vol.2, pp. 799–803.
[9] Miya, K., M. Watanabe, M. Hayashi, T. Kitade, O. Kato, K. Homma, "CDMA/TDD Cellular Systems for the 3rd Generation Mobile Communication," *Proceedings of 47th Vehicular Technology Conference (VTC'97)*, May 1997, Phoenix, Arizona, USA, Vol. 2, pp. 820–824.
[10] Povey, G., H. Holma, and A. Toskala, "TDD-CDMA Extension to FDD-CDMA Based Third Generation Cellular System," *Proceedings of ICUPC'97*, San Diego, California, USA, October 1997, pp. 813–817.
[11] Omura, J., and P. Yang, "Spread Spectrum S-CDMA for Personal Communication Services," *Proceedings of Milcom'92*, October 1992, San Diego, California, USA, pp. 269–273.
[12] Omura, J., "Spread Spectrum Radios For Personal Communication Services," IEEE ISSSTA'92, Yokohama, Japan, November/December 1992. pp. 207–211.
[13] Simon, M., J. Omura, R. Scholtz, and B. Levitt, *Spread Spectrum Communications Handbook*, New York: McGraw-Hill, 1994.

Chapter 10

IMPLEMENTATION ASPECTS

10.1 INTRODUCTION

This chapter concentrates on the implementation aspects of a wideband CDMA transceiver. The goal is to help the reader understand the basic analog and digital structure of a wideband CDMA terminal and get an idea of the problems encountered in real implementation.

Section 10.2 highlights different optimization criteria (power consumption, cost, and size) that should be taken into account while designing a CDMA transceiver. The task is certainly not easy because the number of degrees of freedom is large. The main concern is on the mobile terminal side, but in the future power consumption, cost, and size will become important also for base stations.

In Section 10.3, the transceiver modularity concept is discussed in conjunction with wideband CDMA systems. A wide set of defined third generation system services calls for terminals capable of adapting to different data rates and quality of services. In practice this makes the physical terminal design much more complicated because many functions have multiple sets of operating parameters.

Baseband signal processing is investigated in Section 10.4. The emphasis is on the basic building blocks, but attention is also paid to more advanced techniques like interference cancellation. First, the different processing units are considered from the functional architecture point of view. This is followed by a discussion of how to make the mapping from the functional description to a physical implementation. The problem is not, however, unambiguous, making an optimal solution difficult to find. In all cases, digital application specific integrated circuit (ASIC) or digital signal processing (DSP) -based implementation is assumed although some baseband processing like matched filter could be performed in an analog domain. At the end of the section, different control loops for AGC, AFC, and power control are highlighted.

Section 10.5 presents well-known RF architectures suitable for a wideband CDMA transceiver. The RF section is considered to incorporate everything from radio

frequency signal to baseband, independent of analog or digital implementation. Complexity of the front-end is increased compared to second generation systems mainly because of the need for higher linearity, better selectivity, and larger dynamic range requirements. Most radio frequency processing is still done in the analog domain, while in the future the digital border is moving towards the antenna.

Section 10.6 describes the concept of software configurable radio and section 10.7 illustrates an example configuration of a wideband CDMA mobile terminal architecture, integrating the different parts.

Section 10.8 discusses implementation of multimode terminals. The different implementation requirements between spread and nonspread, between slotted and continuous, between circuit and packet switched, and finally between FDD- and TDD-based systems are discussed.

10.2 OPTIMIZATION CRITERIA

In the design of a wideband CDMA terminal and base station, critical factors to be optimized are power consumption, cost, and size. Each of these can be evaluated by qualitative criteria such as input power requirements, number of components, number of instructions, or ASIC gate primitives. Figure 10.1 illustrates a generic wideband CDMA transceiver partitioned into RF, IF, and baseband sections. Traditionally, RF and IF sections have been realized by analog technology, and baseband by digital technology. However, the interface between analog and digital sections depends on the selected receiver and transmitter architectures, as discussed in Section 10.5.

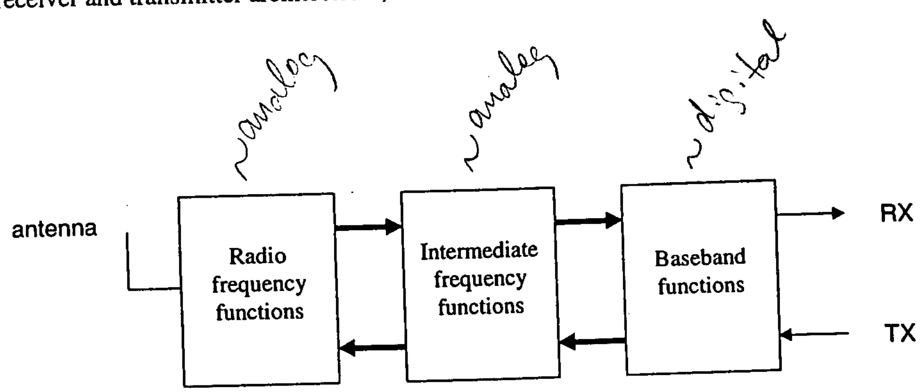


Figure 10.1 Block diagram of a CDMA transceiver.

10.2.1 Power Consumption, Cost, and Size

Due to the radical evolution in the digital signal processing field in terms of DSP processors and new ASIC technologies (see Table 10.1), the analog RF section is becoming the major power consumer in normal speech terminals. On the other hand, more complex baseband functions are being required (e.g., multiuser detector), and if the digital-analog border shifts towards antenna (see Section 10.5.2), the baseband may still form a significant part of the power consumption.

Table 10.1
Development of DSP and ASIC Technologies

	1992	1994	1996	1998	2000
Gate density (gates/mm^2)	2000	5000	15,000	30,000	50,000
Processing power (MIPS)	20	50	80	120	200
Relative power consumption (1/MHz)	100	30	10	3	1
Relative power consumption (max MIPS)	100	75	40	18	10

Source: [1].

On the RF side, the dominating current consumers are the power amplifier and frequency synthesizers. The former is seen as a critical component in the third generation systems due to high linearity requirements likely causing a loss of efficiency. It should, however, be noted that due to smaller cell sizes (pico/microcells), output power is reduced, and hence, the overall power consumption can be kept acceptable. The power consumption of a synthesizer is directly related to the activity factor and is thus larger for systems such as CDMA, which transmit continuously. In addition, two synthesizers are needed in CDMA terminals since transmission and reception take place simultaneously in contrast to TDMA terminals, which need only one synthesizer.

The transceiver cost factor depends quite evenly on every required component. So far, the division between analog RF and digital baseband sections has been approximately fifty-fifty. However, the trend leans toward the analog side because the integrated digital processing side is shrinking into a single chip.

The actual analog and digital data processing functions (encoding/decoding, filtering, mixing, equalization) are not very critical in terms of the total terminal size. This may have been the case in the past, but today the dominant components are the battery and the user/external interfaces. On the data path side, some RF section components such as the duplex filter and possibly the antenna may, however, occupy a large amount of space.

The development of the first third generation mobile terminals is a difficult task. They should have all the features of the second generation systems and include new features related to higher bit rate services. The fact that consumers also expect to be able to connect to the existing systems calls for a dual/multimode engine.

10.2.2 Evaluation Methods

There are a few quantitative parameters that can be used to evaluate the transceiver power consumption, size, and cost. This section explains what should be taken into account when assessing analog and digital section complexities. Radio frequency parts are seen as analog, and baseband functions as digital. IF processing lies somewhere in between.

On the analog RF side, power consumption can be assessed by estimating the average input power requirements in both talk and idle modes. The power consumption is a function of user data rate, activity factor, required effective radiated power, and transmission spectral mask (linearity requirements). Emphasis is on deriving the power amplifier efficiency in different situations. The activity factor of a synthesizer has a large impact on the power consumption. Therefore, effective power-down/up procedures should be used to turn off the synthesizer whenever possible.

Cost and size of the RF section depends strongly on the number of individual components. The major problems relate to isolation requirements and integration of different semiconductor processes, which prevent a single chip RF front-end implementation. Increased integration, especially on the intermediate frequency side, has already been achieved (possibly joined with some baseband blocks), but still there are a few discrete components such as channel selection and duplex filters that must be implemented externally.

Evaluation on the digital processing side is a much more straightforward task because an analytical approach can be followed. Each function can be divided into instruction and/or gate primitive operations, which can further be mapped to real DSP processors and ASIC processes to give direct estimates for power consumption, cost, and size. The most problematic task here is to compare DSP- and ASIC-based implementations. A good method for doing this is to take the word length requirements into account when counting the MIPS requirements. This leads to a concept called normalized MIPS (nMIPS), which states the number of 1-bit operations needed to execute within a second. The example below compares the complexities of a RAKE receiver correlator and complex phase rotation primitives. The former is a clear ASIC-oriented function, while the latter is more ideal for a DSP processor.

Example. The despreader is for a QPSK chip modulated signal with an input sample rate of 4.096 Msps and input word length of 4 bits. The spreading ratio is 64. The same spreading code is used in both I and Q branches, so only two correlations are needed. The phase rotator takes a complex 10-bit input vector and multiplies that with another 10-bit vector. The operation rate is 64 Ksps. In this case, multiplication is assumed to take (Wlin · WLin) and addition (Wlin + WLin) primitive operations. Thus,

$$\text{nMIPS(despreader)} = 2\cdot Fs\cdot[WLin\cdot WLc+((WLin\cdot WLc)+WLin2)] = 148\ \text{nMIPS}$$

where Fs = sampling frequency = 4.096 Msps, WLin = input sample word length = 4, WLc = reference code word length = 1, and WLin2 = integrator second input word length = 10.

$$\text{nMIPS(rotator)} = Fs\cdot[4\cdot WLin\cdot WLin+2\cdot(WLin+WLin)] = 28\ \text{nMIPS}$$

where Fs = sampling frequency = 64 Ksps and WLin = input sample word length = 10.

10.3 MODULARITY CONSIDERATIONS

Third generation wideband CDMA terminals will provide a set of different services varying from a low rate speech service to a high rate, high quality service for data. When designing RF and baseband architectures this should be taken into account. The more adaptable the architecture, the better the available processing resources can be utilized to provide this variety of operation modes.

In wideband CDMA systems, the higher data rates can be obtained by decreasing the spreading factor or by introducing more parallel code channels (see Section 5.8). The latter option, especially in the case of only a few parallel channels, results in a high peak-to-average power requirement and thus to high backoff. Therefore, it is not very good from the transmitter power amplifier point of view. It may be best to use parallel channels in the downlink to obtain good modularity and variable spreading factor in the uplink to avoid parallel transmission in the mobile station.

The baseband section chip rate elements have a fairly fixed complexity as a function of data rate. This is because the data rate is mainly defined by the correlator dumping period, not the chip/sample rate. Multiple code channels, however, create extra hardware requirements, but still many blocks (like receiver front-end, multipath delay estimation, and RAKE finger allocation, as well as complex channel estimation units) may not need to be duplicated for different codes. The critical parts from the complexity point of view are the transmitter pulse shaping filter, additional RAKE receiver fingers, and code generators. The complexity of a pulse shaping filter depends heavily on the input word length, and in the case of several parallel code channels plus high adjacent channel attenuation requirements (long filter), implementation becomes difficult. Each parallel code channel used requires code generator(s) for spreading and despreading. With short M-sequences (simple shift registers), the complexity increase can be tolerated, while with the long Gold and Kasami codes, which require several shift registers, the silicon area consumption will grow significantly. Although a RAKE finger itself is a reasonably simple device, the total receiver complexity increases because each multipath component requires as many despreaders as there are parallel code channels.

The main effect of increasing user bit rate is, however, seen on the narrowband side because processing requirements at the symbol rate are closely tied to the user data rate. This mainly affects the channel encoding, decoding, and interleaving functions. The key aspect here is how to make the architecture as scalable as on the wideband side.

The required scalability can be achieved in two ways. Extra processing power is achieved by installing several processing units operating in parallel; or existing unit throughputs are increased by using a higher clock rate. The former seems justified from the ASIC technology evolution point of view because the available silicon area is constantly expanding. Clock rate increase can be used to save silicon but on the other

hand, may cause hotspots. However, it is not desirable to have high rate system clocks at the mobile station.

Another thing to consider is the specified set of different coding schemes for several bearer services. It is obvious that data services require better BER than speech services. This must be taken into account while designing the encoder/decoder architecture in order to provide maximum resource utilization. The more common lower level processing primitives that can be found, the higher the resource sharing that can be provided. This is easier with encoders, while decoders tend to have more individual constructions. A good example is to use the same physical shift register for convolutional and Reed-Solomon encoders. Software oriented implementation would be a good solution here, but for high data rates the processing power requirements may be too high and too expensive.

10.4 BASEBAND SECTION

Baseband processing implements all the data path functions operating on a signal whose spectrum center is located at zero frequency. In wideband CDMA, the baseband section can be divided into bit/symbol and chip/sample domains, each having a different operation rate. This section describes the basic baseband functions (data path/control) required for a W-CDMA terminal implementation. In all cases, the physical implementation is assumed to take place in the digital domain (not counting the converters).

10.4.1 Baseband Receiver

Figure 10.2 illustrates the generic block of a direct sequence CDMA receiver. After analog-to-digital (A/D) conversion and filtering, the wideband I/Q sample stream is fed into a correlator bank that performs signal transform (despreading) into the narrowband domain where the multipath diversity combiner collects the channel energy from different RAKE fingers. Channel delay profile estimation is required to resolve the multipaths and to set the RAKE fingers to track a specific code phase. After combining, the bits are deinterleaved and decoded to produce transmitted information bit stream. Multiuser detection (MUD) after the despreading and before the multipath combining can be used to increase the system performance. MUD is primarily base station-oriented function and can be implemented in the wideband domain as well.

10.4.1.1 Analog-Digital Conversion

In a wideband CDMA system, because of the wideband signal, a high sampling rate A/D converter is required. Power consumption and cost are not, however, radically increased because the processing gain obtained by spreading allows usage of lower dynamic range converters than in narrowband systems. The direct sequence signal as such is robust against ADC nonlinearities. For a brief discussion of the aberrations, refer to [2].

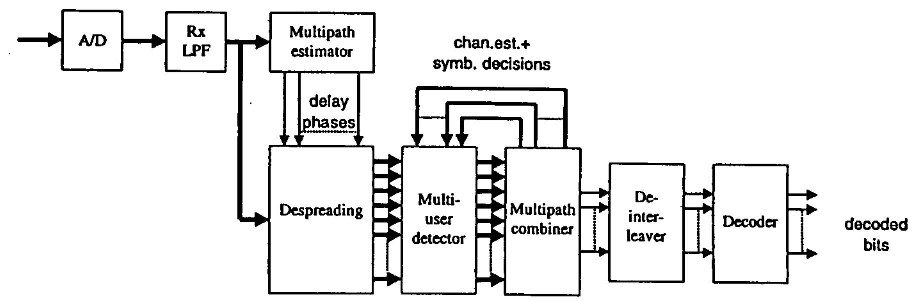


Figure 10.2 Generic receiver block diagram with optional interference cancellation stage.

As stated by Nyquist, the sampling rate must be at least twice the input signal bandwidth. Due to practical reasons the sampling rate should, however, be a multiple of the chip rate (if, for example, 4.096 Mchip/s takes 5-MHz band, it is better to use a rate of 4×4.096 Msamples/s than 2×5 Msamples/s). Furthermore, a higher oversampling rate means that a better multipath resolution can be obtained without utilizing interpolation techniques, which may be quite complex from the RAKE receiver point of view.

The required dynamic range of an A/D converter should be selected in such a way that the converter does not saturate with proper AGC setting and still quantization noise does not contribute to the overall link performance. The CDMA signal differs in this respect from nonspread systems because the actual information bearing signal is buried in noise. For example, with a spreading ratio of 128 and symbol-energy-to-noise density ratio (E_s/N_0) of 8 dB, the carrier-to-interference ratio (CIR) is only −13 dB. In practice, 4-to 6-bit converters have been noted to provide adequate performance [3].

10.4.1.2 Baseband Signal Filtering

An optimal receiver filter is matched to the transmitter pulse shaping filter to provide maximum SNR. A purely digital implementation is too complex; a better choice is to find a compromise between analog IF/BB and digital filtering. In addition, the latter should be designed in such a way that it is able to compensate for the possible signal distortion caused by the transmitter/receiver analog filters and other data path nonlinearities.

The main problem with the receiver filter compared to the transmission filter is the higher input dynamic range (4 to 6 bits) and sampling rate requirement. This means that long filters cannot be provided practically, which is the case especially with the mobile terminal. A shorter nonoptimal filter can, however, be used because it deteriorates the signal quality only slightly. For filter implementation examples in the IS-95 system, refer to [4].

It should still be noted that filtering must be performed to both I and Q signals, which doubles the processing requirements. One option for reducing the hardware complexity is to use only high throughput filters with two delay lines, and clocking this with twice the sampling rate.

10.4.1.3 Signal Despreading

The purpose of signal despreading is to transform the received wideband signal into a narrowband signal. This can easily be implemented using the despreading circuits shown in Figure 10.3. (See Section 5.6.2 for discussion on different spreading circuits.) The number of required correlators depends on the modulation method, possible multi-code reception, and the number of multipath taps to be despread per code channel. The dumping rate of the correlators is set according to the spreading ratio. The wideband clock may be set equal to the chip clock given that the delay phase is correct. This can be achieved by utilizing an interpolator before each RAKE finger, or a higher accuracy clock on the multipath estimator side. In the latter case, the despreader wideband clock is divided from the higher rate clock, and the phase of the chip rate clock must be adjustable.

It should be noted that by using BPSK spreading, the number of correlators can be halved. The case is the same with QPSK spreading if the cross-correlations between the I and Q branch codes are zero. Still another way of reducing the number of correlators is to use the same spreading code for both I and Q signals. This approach is, however, vulnerable to phase errors and requires more accurate channel estimation.

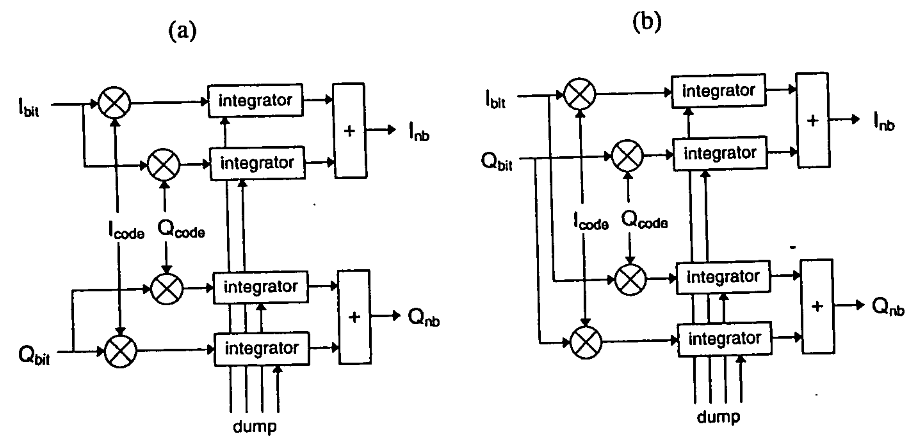


Figure 10.3 Despreader architectures for QPSK chip modulated signal: (a) dual-channel despreading and (b) complex despreading.

Especially in the downlink, flexible data rates can be implemented by using multiple parallel code channels. Extra correlators are needed for each new code, and if the number of codes increases to several tens, mobile terminal implementation becomes cumbersome. Phase and delay estimation can, however, be shared since each code channel passes through the same multipath, and hence, some alleviation can be obtained in terms of implementation.

Higher data rates implemented with the variable spreading factor (VSF) scheme result in a receiver implementation whose hardware complexity is fairly constant as a function of the data rate. Only despreaders for the multipath taps of a single code channel are needed, and the different data rates are implemented by changing the correlators' dumping periods. The integration unit must be designed according to the largest spreading ratio, which defines the worst case register word length.

An important aspect is the number of bits required for the wideband signal presentation. The implementation complexity and cost of the receiver depend heavily on the required wideband signal dynamic range. The input signal consists of additive white Gaussian noise mainly due to the receiver front-end and different user signals propagated through multipath channels. Receiver gain control is used to adjust the input signal level suitable for the A/D-converter dynamic range. By noting that quantization noise depends on the desired input signal level with respect to the converter input range, an estimate for the needed word length can be obtained. Figure 10.4 shows signal-to-noise degradation with respect to the effective quantization word length.

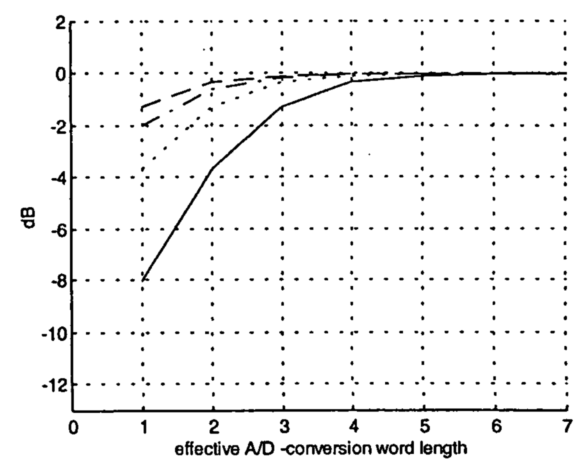


Figure 10.4 SNR degradation as a function of quantization word length. The four curves are with different Gaussian noise variances δ_n into ±1 A/D input (dashed: $\delta_n = 0.5$, dashdot: $\delta_n = 0.375$, dotted: $\delta_n = 0.25$, solid: $\delta_n = 0.125$).

10.4.1.4 *Channel Estimation*

The channel estimation block estimates multipath delays and complex tap coefficients (phase and amplitude). There are numerous different algorithms and approaches to solve the estimation problem. Some of the more frequently used ones are discussed below, emphasizing the actual implementation.

The main target of the multipath delay estimation unit is to provide accurate enough delay estimates at a fast enough rate. Traditionally, the implementation has utilized a sliding correlator that calculates one delay tap power per dumping period (see, for example, [5-6]). The more parallel correlators available, the faster the multipath window can be scanned. A good criteria for finding the "real" taps and rejecting the false ones is to employ so-called two-dwell approach [7]. The procedure has two parts: first, the candidates are found, then a verification is performed. A more exotic approach is to use the subspace concept in which the received signal is projected against the desired user code (vector spanning a subspace) in different delay phases [8]. Should there exist a multipath tap, the projection value will exceed a threshold.

The accuracy of the delay estimate is generally defined by the sampling rate. If, however, interpolation techniques are used, better accuracy can be obtained at the expense of increased complexity. Figure 10.5 shows an example of SNR degradation due to an erroneous delay estimate. It is interesting to note that the filtered input signal is more robust against delay errors that a nonfiltered square wave. The solid line presents correlation with raised root cosine (RRC) filtered reference sequence (input is also RRC filtered); the dotted line is with rectangular input; and the dashed line is with both rectangular input and reference signals. There are four samples per chip in each case.

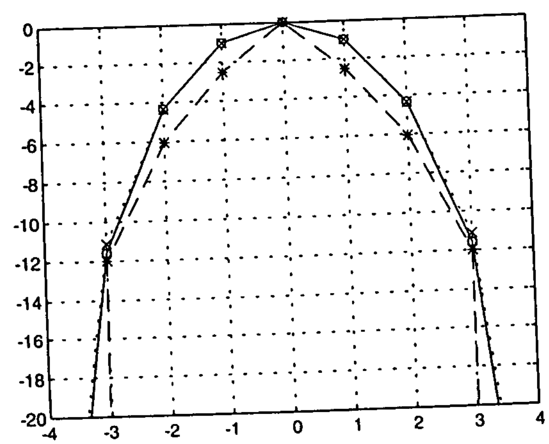


Figure 10.5 SNR after despreading as a function of timing error.

A commonly adopted approach, shown in Figure 10.6, is to use a coarse delay estimation unit, which triggers delay locked loops (DLL) connected to the despreader fingers. The DLL forms a signal proportional to the code phase error, which is used to adjust the code generator(s). The time constant is made long to provide a good enough estimate without excess phase jitter. Implementation of the sliding corrector is simple, but extra complexity is introduced because of decentralized control. In real conditions, multipath taps tend to merge, locking two physical fingers to the same tap. This calls for an indication signal from each RAKE finger to the multipath estimator control, carrying the corresponding finger code phase value.

The system in Figure 10.6 operates as follows. First, the coarse multipath estimator scans the defined channel impulse response window defined by the maximum delay spread. By averaging, the decision logic selects the highest peaks and allocates available RAKE fingers to corresponding code phases. Each RAKE finger's DLL locks to the phase and starts following the drifting multipath tap. From time to time, in addition to a "synch.lost"-indication, RAKE fingers send code phase information to the allocation control, which checks that no peaks are tracked by two or more units. In addition, the multipath control must continue updating the channel impulse profile to detect possible new peaks.

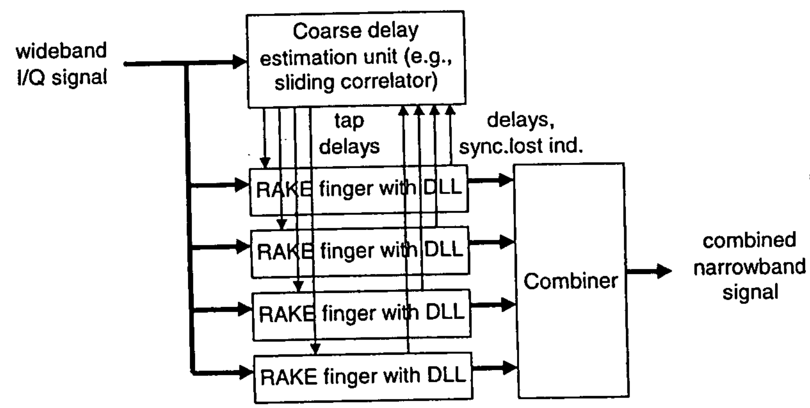


Figure 10.6 Traditional DLL-based receiver.

Today's ASIC technologies are at the point that a full digital code matched filter can be implemented without great effort. In practice this means that the delay estimation unit can be made fast and accurate, suggesting that the DLLs are no longer needed. Impulse response averaging must naturally be applied in order to prevent RAKE allocation due to a noise event. The advantage is a centralized control, which determines

both RAKE finger allocation and the releasing strategy that can be optimized according to changing channel conditions and available hardware (number of fingers).

Complex amplitude estimation (i.e., phase and amplitude estimation) performed for individual channel taps is required for coherent detection. Complex channel estimation is performed with the help of a pilot channel consisting of known transmitted symbols (see Section 5.2.2). The accuracy of the estimate is essential from the link performance point of view, and it depends on the pilot channel energy, the algorithm used, and environment conditions. Variable mobile speed grades require a lot from the algorithm and in the extreme case, call for an adaptive solution.

The pilot channel can be provided in two basic ways. In the continuous pilot case, there is one physical code channel dedicated fully to constant pilot symbol transmission (code multiplexed pilot), while another option is to insert pilot symbols into the data stream (time multiplexed pilot). The former is widely used in the downlink transmission, while the latter approach is needed to implement a coherent uplink. The former argument is not, however, true if adaptive downlink beamforming is applied. Here each MS faces a different channel, which is also the case with uplink transmission.

Figure 10.7 shows one possible phase estimation architecture based on a continuous pilot channel. Each RAKE finger comprises extra correlators for despreading the pilot channel in the same code phase as the data channel tracked by the finger. The obtained instantaneous phase estimate is filtered with a suitable time constant, which can also be made adaptive to the Doppler frequency. Note that before passing the signal to the combiner, a maximal-ratio phase correction is made. The pilot channel correlator can utilize any spreading ratio regardless of the data dumping period.

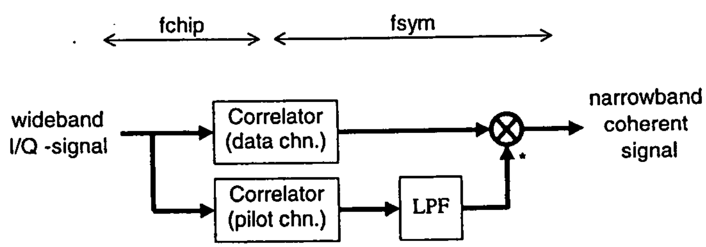


Figure 10.7 Continuous pilot signal-based phase estimation.

In case the pilot signal is transmitted as symbols in conjunction with the data channel, there is no more need for extra pilot correlators and code generators. Two well-known methods of performing the estimation for these are based on either interpolation or decision feedback techniques.

The decision feedback approach shown in Figure 10.8 performs hard decisions for the nonpilot symbols and removes the data modulation. Hence, a continuous pilot

can be constructed and later filtered with a low-pass filter. Pilot symbols within the data are mainly used for taking care of phase ambiguity testing.

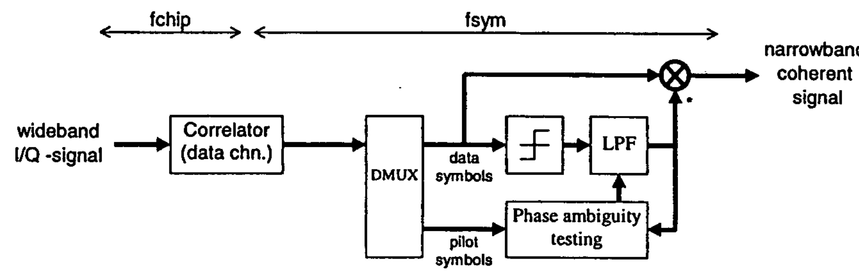


Figure 10.8 Complex phase estimation using decision feedback techniques.

If low rate coding, resulting in very small symbol energy, is used, the decision feedback performance starts to degrade because of increased symbol errors. In this case the interpolation technique may suit better because it uses only the pilot symbol energy for the estimation. The pilot energy can be increased by applying more symbols or by using a higher transmission power. The interpolating system faces difficulties with high mobile speeds because in order to get a good enough estimate a long time constant must be used and thus, the system cannot track the channel. Another problem here is the introduced data path delay due to noncausal estimation. Figure 10.9 shows an example architecture for a RAKE finger with an interpolating estimator.

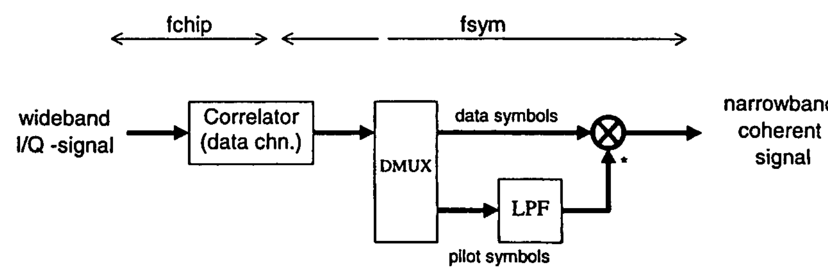


Figure 10.9 RAKE finger with interpolating phase estimator.

10.4.1.5 *Multipath Combiner*

A multipath combiner takes the narrowband outputs from the different RAKE fingers and combines them with each other. Maximal ratio combining produces the best performance. It weights the multipath symbols according to the SNR before combining them.

From the receiver architecture point of view, the tap weighting (maximal ratio case) can be implemented in a straightforward manner inside the RAKE fingers. This is especially the case with coherent detection where each finger incorporates a signal rotation. The rotation is done by multiplying the received signal with the complex conjugate of the phase error vector having length 1. Weighting can easily be merged into this process by setting the length of the vector according to the corresponding multipath tap SNR. In practice this means that the only task for the combiner unit is to add the coherently detected and weighted signals together.

Each RAKE finger does not always process a strong multipath component (e.g., due to fading). If the SNR of a tap is low enough, it is better to leave it out of the combining process. This is because it might have a destructive effect on the total signal quality. Some SNR level value (possibly adaptive) can be used to decide which taps are taken into the combining process.

10.4.1.6 *Multiuser Detection/Interference Cancellation*

The principles of multiuser detector/interference cancellation algorithms were presented in Chapter 5. A large number of different suboptimum multiuser detection approaches exists. This section concentrates on describing the best known suboptimal approaches, namely, decorrelating, MMSE, parallel interference cancellation, serial (successive) interference cancellation, decision feedback detectors, and neural network-based detectors from the implementation point of view.

Interference cancellation is a baseband processing function, which is one of the main advantages over the other capacity enhancement approaches such as adaptive antennas. State-of-the-art digital DSP and ASIC technology already provide an opportunity of building some kind of interference cancellation unit, although optimal implementation (MLSE) is still far too complex.

Decorrelating Detector. A decorrelating detector is depicted in Figure 10.10. A matrix inversion is used for the cross-correlation matrix, and the K signals from the traditional RAKE receiver branches are filtered with this R^{-1} operator. From the implementation point of view, the system becomes complex because the matrix inversion must be calculated every time the users' mutual time delays change. The Cholesky decomposition [9] can be used to perform the matrix inversion faster and hence ease the processing. The matrix is not always invertible (i.e., the matrix is singular) although according to [10] this should not happen in most cases. From the computation point of view, the precision required to perform the inversion is also critical.

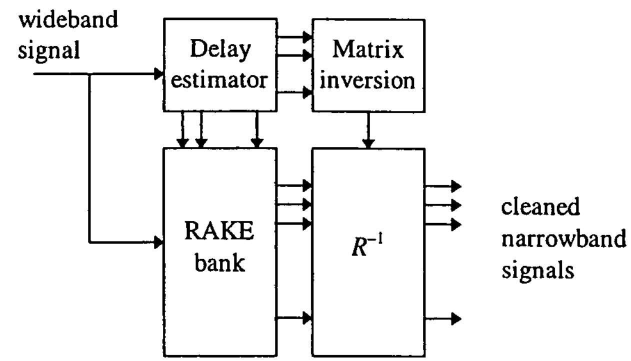


Figure 10.10 Decorrelating detector.

A decorrelating detector can also be implemented in the wideband domain. In this case, the traditional type correlator receiver is used while replacing the reference chip sequence with a signal orthogonal to all the other received users [11]. From the complexity point of view, the difference is the multilevel nature of the despreading signal and the processing element to calculate this sequence. The latter causes a radical complexity increase if operated dynamically, and hence, a suitable application would be a single user detector applied to a synchronous system.

In practice, the implementation complexity of the direct matrix inversion might be too high, especially for a large number of users and multipath components. Therefore, as was discussed in Chapter 5, iterative algorithms such as the conjugate gradient (CG) method have been proposed for implementation of the decorrelator.

Minimum Mean Square Estimate (MMSE) Based IC. From the performance point of view, a decorrelating detector has the disadvantage of noise power enhancement in the matrix inversion process. An MMSE-criteria-based interference cancellation receiver overcomes this by taking the noise power into account. The matched filter output is fed through a linear transform derived according to MMSE criteria [12]. The linear operator is obtained by taking the inversion of the cross-correlation matrix that is conditioned with received user signal powers and estimated noise power on the diagonal. Performance improvement comes, however, with extra detector complexity. This is due to the need to estimate individual user signals and noise power levels. The MMSE-based receiver performance is similar to that of the conventional matched filter with low SNRs, while with high SNRs, the performance approaches the decorrelating detector.

The MMSE receivers have attracted interest because of their applicability to decentralized adaptive implementation. Adaptive implementation does not require information of the interfering spreading codes, and thus, the adaptive MMSE receivers could be used in the mobile station as well. However, they require the signal to be cyclostationary (i.e., periodic). Therefore they can be used only with short spreading codes.

Matrix inversion turns out to be the main source of implementation complexity. Because the matrix is symmetric and still positive definite, Cholesky's method can be used to ease the calculation. It should still be noted that as in the case of the decorrelating detector, the inversion must be done every time the mutual user multipath delays change. In addition, the update must be performed every time the user or noise signal powers vary. Hence, similar to the decorrelator, iterative algorithms should be used.

Parallel Interference Cancellation. Parallel interference cancellation (PIC) cancels the interference of all users simultaneously. The multistage PIC shown in Figure 10.11 suppresses the multiple access interference in multiple consecutive steps. Stage n cancels the interference by utilizing the hard symbol decision from stage n - 1. The cancellation process can be performed either in narrowband (NB) or wideband (WB) domains. For the corresponding detectors we use notations NB-PIC and WB-PIC. The cancellation operation itself requires knowledge of the user codes, the relative code phases, complex channel estimates for the multipath taps, and finally hard symbol decisions for each user.

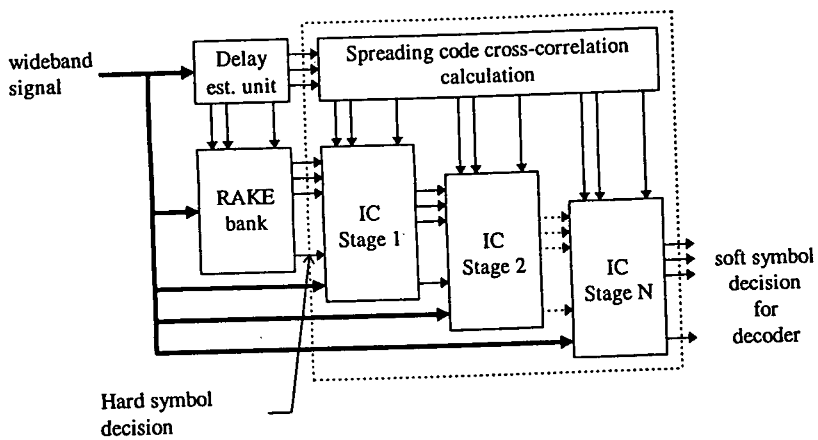


Figure 10.11 Multistage PIC detector with two cancellation stages.

Multistage PIC [13] operating in the narrowband domain performs computation with a symbol rate. Hard symbol decisions from the preceding stage are used in conjunction with complex channel estimates and mutual normalized cross-correlations between the spreading codes to generate aggregate interference estimates for each despread multipath component of interest. Multiple access interference (MAI) cleaned multipath signals are combined for each user, and more accurate hard symbol estimates are made from these. Most of the gain can already be obtained with two cancellation stages [13].

From an architecture point of view, implementation of the NB-PIC detector is a straightforward procedure because it can be fitted fairly directly into a traditional CDMA receiver between the correlators and the deinterleaver/decoder. Multipath combining and hard symbol decisions are inherently included in the interference cancellation stages. The actual cancellation operation is simple because it takes place at the symbol rate. The total processing requirements become high, however, for two reasons. First, the spreading code cross-correlations must be updated every time the mutual tap delays change. In the extreme case, if long spreading codes are used, the whole matrix must be recomputed for every symbol. In addition, the aggregate interference estimates depend on the cross-correlations and differ for each multipath. Hence, we need to compute at symbol rate $K \times L$ interference estimates each consisting of ($K \times L$ - 1) additions and multiplications (K = no. of users, L = no. of multipath taps per user).

Physical implementation of the NB-PIC detector depends on the symbol rate, spreading factor, type of spreading, number of stages, and the total number of multipaths involved in the IC process. The interference estimation is likely to need an ASIC-based solution, while the actual cancellation could be done with DSPs. Multiple stages can be implemented using either serial computation with one physical instance or if the processing rate is too high, with parallel instances in pipeline mode.

Should long spreading codes be used, the receiver complexity due to the constant need for cross-correlation updating is increased vastly. A wideband version of PIC, also called regenerative PIC, would be a better choice in this case. Finally, it should still be perceived that the narrowband multistage detector architecture relies strongly on unified code channel symbol rate, making it more suitable for systems utilizing parallel code channels instead of variable spreading to provide variable bearer services.

A major problem with the NB-PIC detector is that each multipath component, due to different mutual cross-correlations, sees a different interfering signal on the narrowband side. This can be prevented by conducting the cancellation in the wideband domain. WB-PIC, shown in Figure 10.12, is based on the regeneration principle, which is used to calculate the so-called residual signals. Regeneration, which requires knowledge of complex channels, hard symbols decisions, tap delay estimates, and the spreading code, is done for every multipath component taken into the cancellation process. The combined regenerated signal estimate is subtracted from the received wideband signal to get the residual signal. This is added to the regenerated individual multipath components in order to get a cleaned wideband signal for the despreaders.

WB-PIC has the advantage that it can be utilized with both short and long code based systems. In addition, the residual signal provides an interference resistant input for multipath profile estimator. Although the actual interference cancellation operation requires a higher operation rate than in the NB-PIC case, the lack of cross-correlation calculation makes the total processing/resource requirements smaller. The complexity of the regeneration part depends heavily on the pulse shaping filtering scheme used and hence on the sampling frequency in conjunction with the word length. The type and length of the spreading codes also have a noticeable effect on the total system complexity.

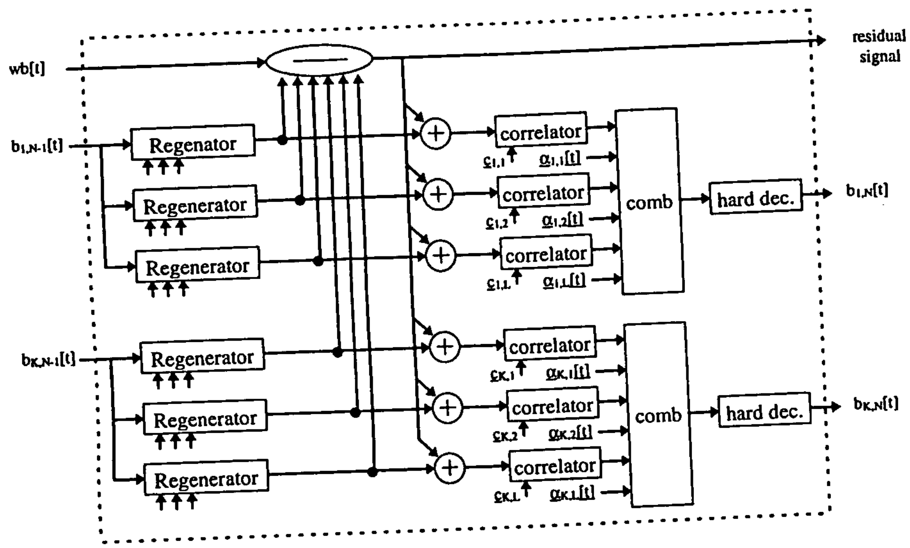


Figure 10.12 Wideband domain parallel interference cancellation (WB-PIC) stage.

Multiple simple parallel operations in WB-PIC functional architecture suggest clearly an ASIC implementation. A fully parallel architecture with multiple physical stage instances provides the lowest latency, while one WB-PIC operated sequentially would require less gate area. It would also be easy to perform serial userwise, or even pathwise, processing to further minimize the area. WB-PIC indeed suits many different processing schemes, which makes it a potential candidate for a real application.

In the previous text concerning multistage detectors it has been assumed that the interference cancellation comprised all the users with their multipath components. In practice the receiver could be made adaptive in such a way that only the strongest users/multipaths are canceled. This would require extra system/control resources for decision making but otherwise would reduce the actual PIC complexity. The system performance may, however, degrade slightly due to the selective cancellation, but in some cases it may even increase. The latter is the case when weak users are left out from the IC process. The above idea seems like a good one, but we must remember that in DS-CDMA the power control tries to keep the receiver signal levels alike. On the other hand, this is not actually the case when the VSF scheme is used for higher bit rates or when signal power control is used to adjust the quality of service.

Successive Interference Cancellation [14]. Successive interference cancellation (SIC), shown in Figure 10.13, ranks the users according to their power levels. First, the user with the strongest signal is detected, and the MAI estimate for the following users is updated with this knowledge. Next, the second strongest user is processed continuing to the weakest one. From the implementation, and also from the performance point of view, the process can be stopped after cancellation of a subset of users (the most interfering ones). The cancellation can be done on the wideband or narrowband side. The former uses the canceled user symbol hard decision, signature waveform, and amplitude and phase estimates to regenerate the wideband signal. This is then subtracted from the received wideband input stream, which is later despread by the rest of the users. Another method is to use cross-correlations between the spreading codes of the strong user and of the "cleaned" users, and do the cancellation after the despreading operation. It should be noted that the cleaned wideband canceled signal can be used in the multipath channel estimator, thus making it MAI resistant.

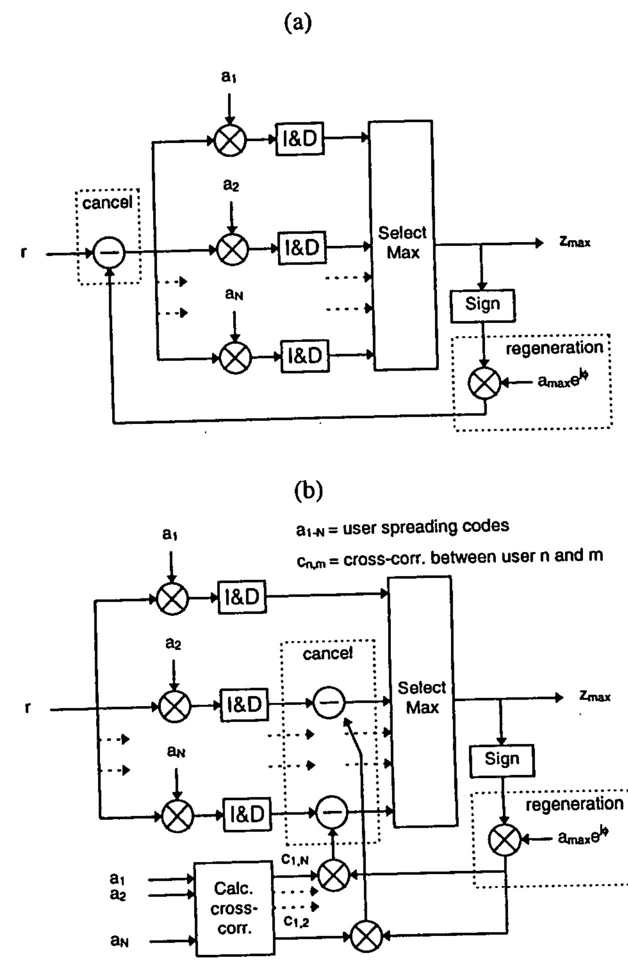


Figure 10.13 Wideband- and narrowband- based successive cancellation receivers (*From*: [15]).

Decision Feedback (DF) Detectors. DF detectors combine linear and successive interference cancellation. They take hard symbol decisions for all, or a subset of active users, and feed this value back to estimate the MAI. The receiver generally consists of a feed-forward cancellation part relying on the previous stage symbol decisions, and a feedback part utilizing the following stage "cleaned" hard symbols in estimating the MAI [12]. The times of signal arrivals play an important role here by deriving the order of cancellation. A conventional detector is used as the first stage. The performance can, however, be improved by introducing an extra feedback part already to the first stage [16]. In case of severe near-far reception, a good approach is to rank the users according to the power levels and perform the intra-stage cancellation starting from the strongest user [17].

The decision feedback detectors are very similar to the already-discussed multi-stage ones. With respect to the former, the main difference is that within a single stage, the users are detected sequentially in time (according to power levels, or time of arrival), and hence the hard decisions based on readily "cleaned" user symbols can be utilized with subsequent detection of the same stage users (see Figure 10.14). This could also be called a serial-parallel interference cancellation.

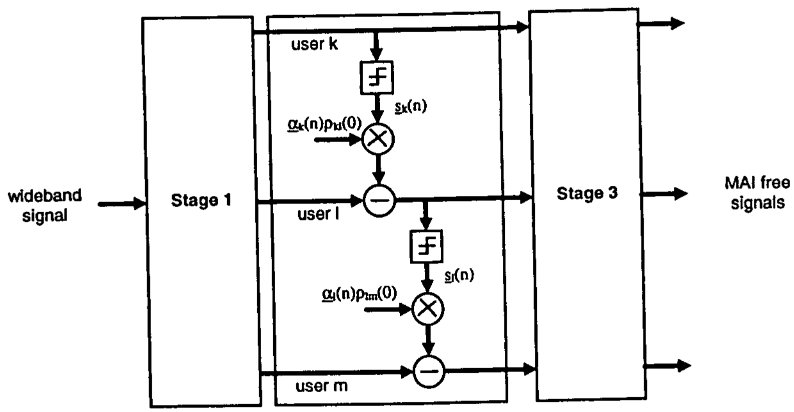


Figure 10.14 Decision feedback detector with three synchronous users.

Neural Network-Based Detectors. Neural networks (NN) provide a means of solving optimization problems with computational elements called nodes that are connected to each other with weighted (possibly adapted) interconnections. The effectiveness of neural networks is due to parallel computation taking place in numerous simple processing elements. In multiuser detection, neural nets can be utilized to find the most likely transmitted bit sequence that produced the matched filter output. There are many different neural network architectures [18], while the most commonly used for MUD realization are the normal feed-forward multilayer perceptron [19] and a simpler Hopfield net shown in Figure 10.15 [20].

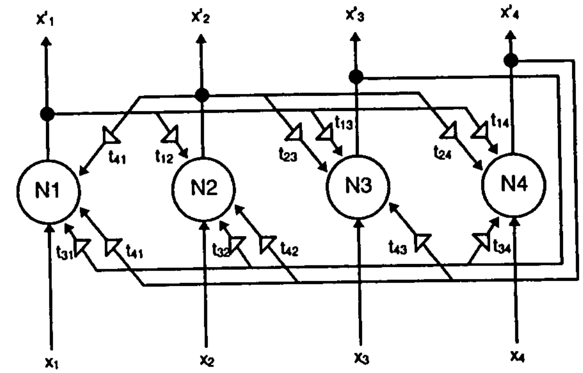


Figure 10.15 Hopfield neural net with four nodes.

A major disadvantage of the feed-forward structure is the prerequisite for training. In addition, the number of nodes increases exponentially with respect to users to be detected. Because of these drawbacks of the feed-forward structure, the Hopfield network is seen as a better neural network choice for multiuser detection problems. Its complexity expands with the square of the number of users, which will also lead to a large system in case of many active users. Training is not needed, while the processing burden comes from the iteration rounds (depends on the number of users). Therefore, neural networks cannot be used to process the whole set of received users, and the number of inputs should be reduced. Figure 10.16 shows an approach utilizing a reduced detector (RD) in conjunction with Hopfield net. The reduced detector implemented an iterative algorithm focusing only on users with nonconfident parameters (reduced search space) [21]. If the iteration turned out to be irreducible, a Hopfield net was used to perform a more extensive search.

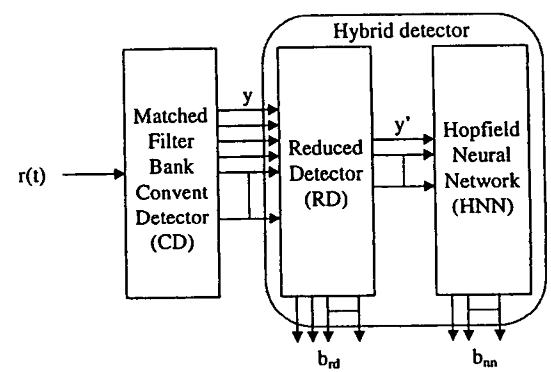


Figure 10.16 Hybrid matched filter, reduced receiver, Hopfield neural net detector (*From*: [21]).

From the implementation point of view, the neural networks are naturally ASIC oriented (many parallel elements). In the past, analog techniques were applied to provide topologies with up to 50 neurons. Digital approaches tend to be problematic for larger nets with nonbinary inputs and interconnections, although this is about to change due to even higher integration levels. Multiplication, addition, and some nonlinear functions are the primitive operations to be provided. The number of multiplications depends on the interconnections, and when binary inputs are used, these are fairly simple. Additions, on the other hand, cause major problem here because the required adder trees call for primitive adders with nonbinary inputs.

10.4.1.7 *Decoding Architectures*

Forward error correction (FEC) and detection coding are used to protect the transmitted information bits and provide the specified quality of service. The most frequently used FEC schemes are based on convolutional and block coding. The Viterbi algorithm is used to perform convolutional decoding, while simple shift registers can be applied for block codes. Third generation cellular systems create many new challenges for the terminal decoders. These are due to the increased number of different bearer services with different data rates, BERs, and delay requirements.

The different operation modes and coding schemes call for modular and flexible decoder architectures. For pure speech terminals the physical implementation can rely on DSPs, while the higher data rates certainly need an ASIC accelerator. The key goal here is to realize which primitive operations are common to the different coding schemes and how these can be effectively utilized to provide variable data rates.

Block coding schemes in conjunction with error detection coding (CRC) can easily make use of the same physical processing resources. This is because the decoding can be implemented using shift register chains, and the different schemes can be supported with variable polynomial lengths and changeable polynomial connections. Constant processing clock rate with alternating activity periods provide a good means for adapting the decoder to different data rates. The modest complexity of the shift register depends naturally on the polynomial degree.

The Viterbi decoder is built on top of a trellis tree consisting of stages and transitions. The basic operation consists of partial metric calculation, trellis updating based on path selection, and back-tracing. The partial metric processing involves calculation of 2^K values (K = constraint length) for each received bit. The total number of the stages is $2^{(K-1)\times L}$ where L is the so-called truncation length. After L bits have been updated to the tree, back-tracing is performed providing the decoded bits. From the complexity point of view, K determines the required processing rate, and, incorporated with L, the total decoder memory requirements can be derived. The truncation length also defines the decoding delay. Based on simulation results, L equal to $5\times K$ is an adequate selection [22]. So far, constraint lengths less than or equal to 7 have been widely used. It is likely, however, that the third generation cellular standards include specification for larger constraint lengths (e.g., nine). This, in conjunction with high bit

rates, means that decoder processing requirements increase remarkably compared to second generation systems.

From the physical architecture point of view, a key primitive with the Viterbi decoder is the so-called add-compare-select (ACS) primitive that is used to select the partial metrics that have lower values. For fast decoding requirements an ASIC implementation is a necessity, while low rate systems may incorporate a software solution (assuming that the selected DSP can provide an ACS command). The best Viterbi decoders at the moment are capable of providing throughputs of 1 Gbps [23]. Sub-optimal Viterbi decoders are mainly based on reducing the number of states in the tree. M- and t -algorithms are examples of these [24]. The latter could also be classified as an adaptive decoder because the number of states is variable. For a comprehensive description of the Viterbi algorithm see, for example [25,26].

Finally, it should be noted that new decoding approaches based on iterative algorithms have emerged. These are referred to as Turbo decoding, and they have claimed to possess better performance compared to convolutional schemes. On the other hand, their main disadvantage is an increased decoder complexity.

10.4.2 Baseband Transmitter

The transmitter baseband section is mainly responsible for channel coding, including both error detection and protection functions. In the case of third generation systems, several different schemes are required to fulfill many service qualities and user data rates. Encoding functions are considered as bit level procedures.

Signal spreading specific to CDMA systems takes place on the baseband side. This operation causes the input bandwidth to widen according to a defined spreading factor or processing gain. Signal spreading is purely chip-level processing.

Digital filtering is commonly used to shape the transmitted signal spectrum. The filtering must be performed for a wideband signal and becomes complex due to higher clock rates. There exist, however, good architectures providing implementation-efficient solutions.

10.4.2.1 *Channel Encoding*

Error correction algorithms can be divided into block- and tree-based schemes, convolutional coding being the best known from the latter category [27]. Recursive convolutional codes, also called Turbo codes, have been under intense research recently.

Traditionally, the implementation of channel encoders has concentrated on finding good solutions for some constant set of parameters like input data rate or coding parameters. This is not the case anymore due to diverse requirements and the new emphasis on flexibility. The encoders for the third generation systems must be capable of operating with different data rates and coding parameters. In the far extreme, the same physical encoder hardware should even support several coding algorithms. Additional problems still arise since the coding scheme might have to be switched on the fly, the requirement of which is due to packet access and link adaptation.

In the second generation systems, to attain flexibility channel encoding has been normally implemented in software mainly because the data rates have been reasonable low. Third generation systems will place much tougher requirements on the encoders due to radically increased user data rates up to 2 Mbps. Fortunately, the well-known encoding algorithms can be efficiently implemented in ASICs. However, ASIC implementation has the problem of not providing high flexibility, but we should remember that most of the encoding algorithms can be implemented using simple shift register chains. This is also the case with Turbo encoders [5].

Hardware encoder throughput can be adjusted by three different approaches (see Figure 10.17). The first option is to provide parallel hardware, and activate necessary resources when needed. This suits well systems with data rate expansion using parallel "low-rate" code channels. In the other options the same physical encoder core is used, but the operating cycle, or the clock rate, is made variable. The latter may be problematic from a mobile terminal point of view.

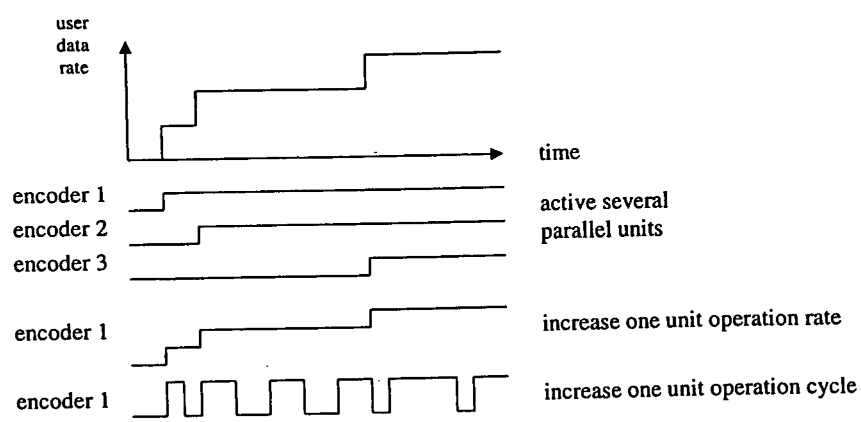


Figure 10.17 Different options to increase the encoder throughput.

10.4.2.2 Data Modulation and Spreading

After channel coding, data modulation is performed. Usually coherent phase shift keying (BPSK or QPSK) data modulation, which provides good spectral efficiency, is applied. The physical implementation of the bits-to-symbol mapping is a very simple operation.

A signal with a wideband spectrum is obtained by multiplying the transmitted symbols by a faster rate chip stream. The operation is highly hardware oriented due to its simplicity (binary multiplication) and its high frequency. Depending on the modulation method, the spreader may include only in-phase or both quadrature branches. In addition, the spreading method defines how the chip sequence(s) is

multiplied with the symbol stream. As shown in Figure 10.18, for an I/Q-signal this can be done individually for both branches (dual-channel spreading) or by using a complex multiplication. The latter option means an increase in processing requirements due to extra multipliers and adders.

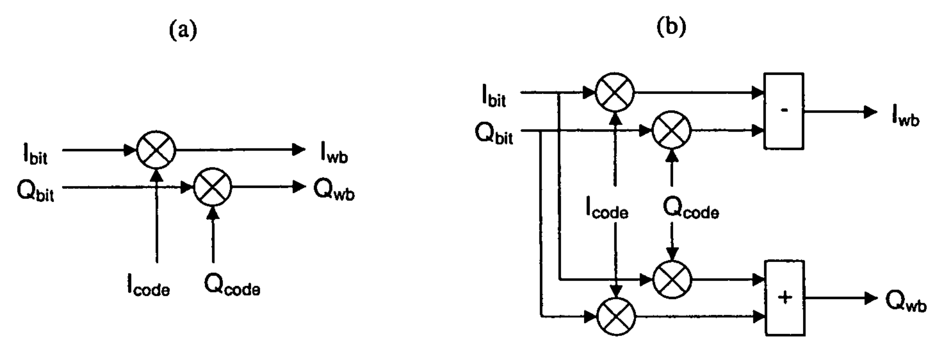


Figure 10.18 (a) Dual-channel and (b) complex spreader architectures.

From the overall transceiver point of view, the spreader procedure has only a minor effect on the total complexity. Some increase, however, may be seen in the baseband power consumption and silicon area, mainly due to fairly long code generators.

10.4.2.3 Baseband Pulse Shaping Filtering

Due to bandwidth-limited output signals, the output chip stream from the spreading modulator is filtered using either a digital or analog filter (pulse shaping filter). A digital filter with linear phase can be designed more accurately but will require a larger dynamic range from the DACs following it. Still, the implementation of long filters (close to 100 taps) may turn out to be too complex. Square RRC-type impulse response is generally used due to the Nyquist no ISI sampling criterion. The roll-off factor defines the sharpness of the spectrum.

For digital implementation, the filter length may have to be truncated in order to minimize complexity. Radical truncation effects on the spectral shape can be alleviated to a certain degree by using a windowing function [28]. Hamming, Hanning, Bartlett, and Kaiser functions are examples of the best known. The filtered output signal must have a high enough word length to keep the quantization noise due to D/A conversion

below spectral mask requirements. As a rule of thumb, one extra bit increases the signal-to-quantization noise ratio by 6 dB [28].

The oversampling factor, which is also an important design parameter, defines the sample-and-hold (S/H) boxcar "filter" and, more importantly, the reconstruction filter requirements. The higher the oversampling factor is, the further apart the copied spectral components appear, which relaxes the post-DAC filter specification. Analog filters, like Bessel or Chebyshev, with two to five taps would be suitable in many cases. Care must be taken, however, to prevent the filter phase response from distorting the signal. Mixed-mode simulations must be performed if tight transmitter specifications are to be fulfilled.

The two plots in Figure 10.19 show the effect of D/A conversion following pulse shaping. The RRC filter with roll-off 0.35 was employed. The signal spectrum is copied into frequencies, which are multiples of the sampling rate.

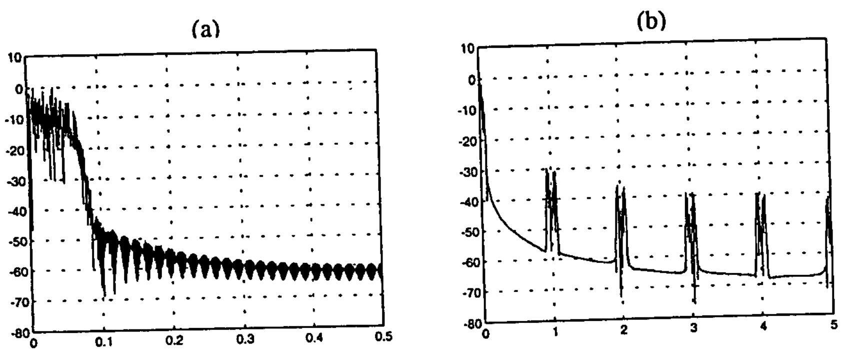


Figure 10.19 Effect of D/A conversion box-car filter on the signal spectrum: (a) signal spectrum after RRC filter and (b) signal spectrum after D/A-conversion.

10.4.2.4 AFC, AGC, and Power Control

In order to make the transceiver work, there are three important functions still to be performed. Automatic frequency control (AFC) is needed in a mobile terminal to synchronize its local oscillator to the base station oscillator. Automatic gain control (AGC) on the receiver side is responsible for preventing the ADC from saturating, while at the same time trying to provide a maximum available dynamic range. The transmitter side still incorporates a power control (PC) adjustment, which is particularly important in CDMA systems, as discussed in Sections 5.11.6 and 7.3. Figure 10.20 depicts the control loops. The dotted lines represent optional connections. For example, in the case of AFC, the carrier frequency can be changed by adjusting the reference oscillator or possibly the digital synthesizers on both the TX and RX sides.

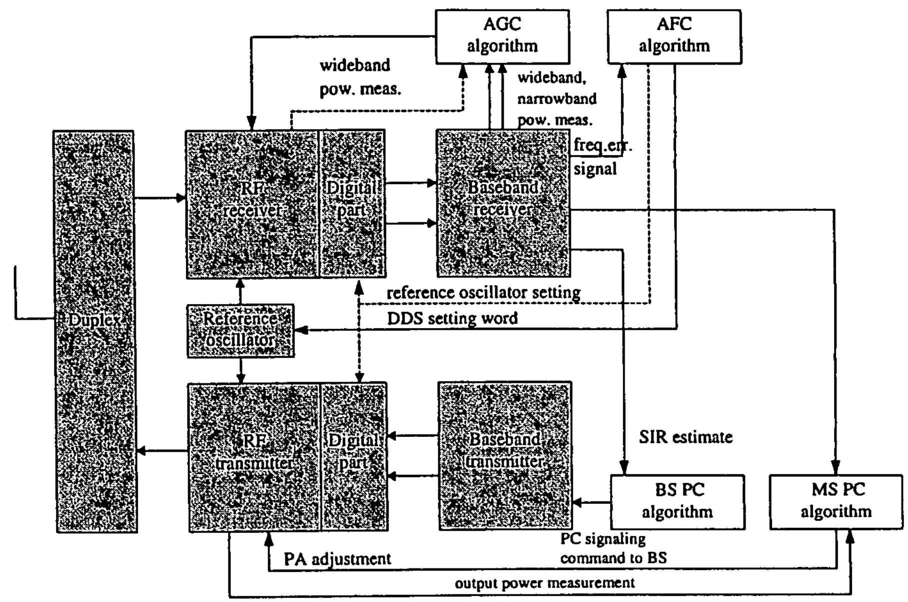


Figure 10.20 Mobile station internal control loops (optional ones with dotted lines).

Automatic frequency control aims at bringing the mobile up/down-conversion frequencies as close as possible to the base station ones. This can be done either by adjusting the MS reference oscillator or, if utilized, the digital frequency down-conversion. Generally, the system digital clocks are also tied to the reference clock, and by performing AFC with the reference, the side effect also provides more synchronized sample/chip/bit clocks. The AFC loop can be made very slow because the oscillator drifting in time is also a slow time varying process. Frequency control could also be performed to compensate for the Doppler effect. This is, however, a slightly different problem compared to slow AFC because on the receiver side the compensation must be done in the opposite direction, which requires a fast absolute frequency error estimate.

The received signal envelope has a large dynamic range. Gain control is required to adjust the amplitude suitable for the A/D converter input range, which tries to provide the highest dynamic range. Finding the optimal operation point is not an easy task because it depends on many aspects like modulation, fading, multipath channel, and, in DS-CDMA systems, on other user signals. AGC loop adjustment is based on the received signal power (RSSI), which can be measured from both wideband and despread narrowband signals. The wideband power can be measured either directly from the RF unit or after the receiver A/D converter.

CDMA systems are vulnerable to erroneous transmission power control because any error will introduce extra cell interference and hence reduce the capacity. The power control loop must be fast enough to react to abrupt signal changes and accurate enough

not to generate undesired power on the channel. The criterion for adjusting the power level should be based on needed service quality; SIR is a good option for this purpose. In the open loop approach, the transmitted power is adjusted according to received signal; while in the closed loop power control, the commands are transmitted as part of the signaling information (e.g., every 5 ms with 1 dB up/down command). In the latter case, the time between the RX power measurement (by BS) and the actual TX power adjustment (by MS) must be kept as brief as possible. In a real implementation the power amplifier's varying characteristics in time should be taken into account. From time to time some power measurement value from the PA output should be used to calibrate the transmission.

10.5 RF SECTION

This section reviews a few of the best known RF transmitter and receiver architectures. In a wideband CDMA transceiver the RF must be designed to suit a wideband low power spectral density signal. Furthermore, transmission is continuous unlike in the case of time division systems. Requirements for high dynamic range, accurate fast power control loop cause most of the problems compared to second generation systems. In addition, the linear modulation and multicode transmission (multilevel signal) place challenges on the power amplifier linearity and efficiency design.

10.5.1 Linearity and Power Consumption Considerations

Third generation systems place much more difficult linearity and efficiency requirements for the RF front-end. The linearity constraint is due to tighter output spectral mask specification, higher signal envelope variations (linear modulation), and, in the case of the PA, the need to keep the operation level near the compression point in order to achieve a high enough efficiency. In addition, when multicode transmission is applied, more backoff is needed causing a loss of efficiency. Adjacent channel interference (ACI) levels of –30 dB are likely to be specified, while with large umbrella cells and flexible channel allocation the requirements may increase to –60 dB.

The transmitted signal spectrum at the antenna defines how much energy is emitted to the adjacent and alternate channels (by alternate channels we mean those after the adjacent ones). Interference is increased due to the spill-over of the signal into adjacent channels, and the network capacity decreases accordingly. As was discussed in Chapter 8, systems incorporating hierarchical cell layers become very difficult because without highly orthogonal spectra, one umbrella-cell user may entirely block microcell traffic due to its radically higher transmission power. Output transmission spectrum depends mainly on the modulation method, multirate solution (multicode vs. variable spreading), transmission filtering, and RF nonlinearities. In the last case, the power amplifier has the largest contribution to the signal distortion.

It should be noted that the overall transmitter linearity is an optimization problem between PA gain characteristics and transmission filtering. The former makes the spectrum narrower but at the same time increases the signal crest factor (= peak-to-

mean signal ratio). The higher the crest factor, the more the signal is distorted by the power amplifier. However, it should be noted that examining only the crest factor may be misleading. From the generated adjacent channel interference point of view, a high crest factor is tolerated if the occurrence of peak amplitudes is low. Hence, instead of crest factor, the power amplifier input signal power histogram should be investigated, as seen in Figure 10.21.

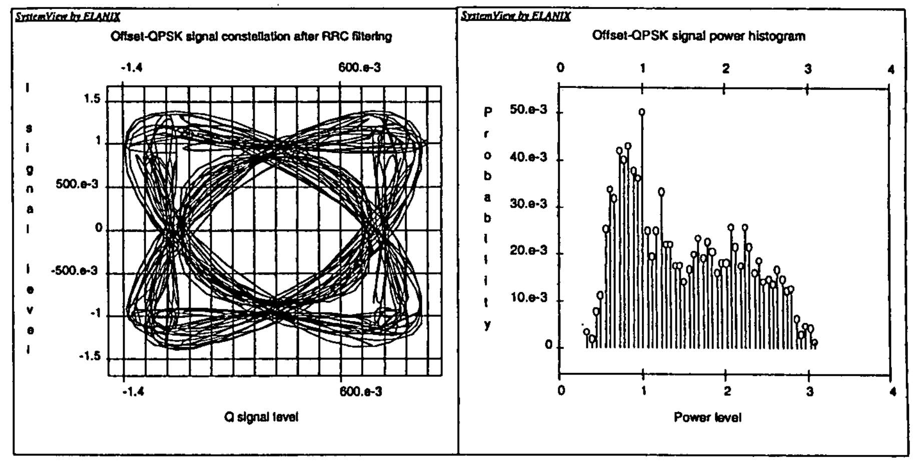


Figure 10.21 Constellation and histogram plots for RRC-filtered O-QPSK signal (roll-off = 0.35).

Power amplifier linearity depends on the operating class. The class can be selected by using different biasing levels. A-class amplifiers are linear at the expense of low power efficiency because they consume power regardless of the input signal level. C-class amplifiers, on the other hand, are power efficient with the introduction of higher nonlinear characteristics. PA linearity can be defined by input amplitude versus output amplitude (AM-AM), and input amplitude versus output phase (AM-PM) relationships. The most nonlinear operation takes place within low/high input signal regions. These characteristics still vary within time due to aging and supply voltage, output load, and temperature change.

The PA can be made linear by applying higher biasing (to A/AB-class), by increasing input signal backoff, or by using linearization techniques. The two former solutions are not very suitable for handheld terminals because the efficiency is decreased accordingly. Linearization techniques are thus seen as the key solution to overcome the tightened spectral mask requirements in conjunction with acceptable amplifier efficiency. It should be noted here that the linearization techniques do not ease the multicode transmission, which will in any case call for extra backoff.

Linearization techniques can be divided into four main categories: (1) feed-forward, (2) feedback, (3) envelope elimination and restoration, and (4) predistortion

[29,30]. Each of these have a set of variants providing different implementation complexity, ACI improvements, and bandwidth/convergence rates. Figure 10.22 shows an example of PA architecture linearized with the complex gain predistortion technique. This approach provides a fairly simple implementation with high bandwidth/convergence rate [29,31]. The input signal power is used to address the predistortion vector stored in RAM. The predistortion vector is updated according to changing environment parameters. Because the number of entries is fairly small, the adaptation is fast. It is likely, however, that the high dynamic range complex multiplier turns out to be the most critical component in the system.

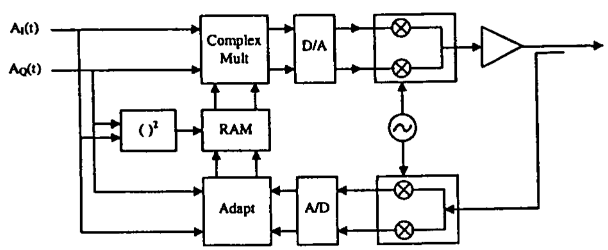


Figure 10.22 Power amplifier linearization with complex gain predistortion.

10.5.2 Receiver Architectures

While the transmitter is facing optimization between linearity and efficiency, the receiver is concerned with providing high selectivity with low noise figure (NF). The non-linearities must be handled in order not to cause excessive signal distortion. Noise figure tells how much the receiver deteriorates the incoming SNR. While more optimal technologies and components have emerged, lower noise figure RX chains can be implemented. Good receivers provide NFs as low as 5 to 7 dB. High selectivity is desired in order to prevent adjacent/alternate channel energy from getting to the input of the ADCs, and decreasing the dynamic range requirements. The former is critical especially when operating in a hierarchical cell system where the nearby channel power may be several tens of decibels higher. A low sensitivity receiver would cause signal saturation in the A/D input.

Linearity of the signal path on the receiver side becomes more important in the case of linear, multilevel modulation or multicode transmission schemes. Still, if the system is operating close to narrowband systems with high power spectral densities, the intermodulation products may distort or even block the whole reception.

The desired channel signal strength may vary for several reasons. The receiver must ensure that if the signal goes up rapidly the ADC is not saturated, or if there is fast power level degradation the signal quality does not pass below an acceptable level. Fast

and high dynamic range (up to 80 dB) AGC is responsible for adjusting the receiver variable gain amplifiers and attenuators in order to feed the ADC with as optimal a signal level as possible. In CDMA systems the adjustment can be done every symbol time without regard to slot boundaries. This is because the spread bits inherently include a training sequence, and channel estimation/equalization takes place on the symbol level.

10.5.2.1 Super-Heterodyne Receiver

The super-heterodyne architecture shown in Figure 10.23 is the most well-known receiver to implement good selectivity. The RF signal is filtered, amplified, and converted to an intermediate frequency in which the channel filtering is performed. The IF frequency is kept constant, while channel selection is done by changing the RF mixer frequency.

Several stages with different IF frequencies are needed to get the desired performance at the expense of higher receiver cost. A single conversion receiver is the simplest option and is commonly utilized.

From the total receiver noise figure point of view, the low noise amplifier (LNA) preceding the RF mixer is the key component. A typical LNA has a noise figure below 4 dB, while the best offer an NF around 1 dB with power gains up to 30 dB.

Good channel selectivity with feasible linearity can be achieved by using surface acoustic wave (SAW) filters at the IF frequency. The size of the filter depends on the channel bandwidth and IF frequency and hence can be implemented smaller for wideband CDMA systems. The problem with the SAW filter is its high insertion loss and relative cost. As a result, solutions to get rid of it have been searched for (direct-conversion and digital receivers being such examples).

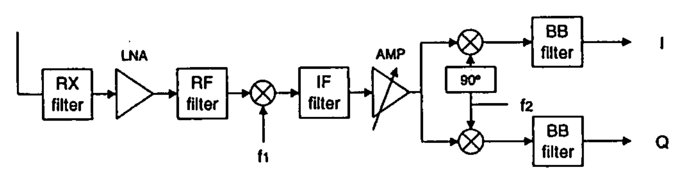


Figure 10.23 1-stage super-heterodyne receiver architecture.

As was the case with the IF-based transmitter, digital direct synthesis (DDS) can be utilized in super-heterodyne receivers to perform I/Q separation (see Figure 10.24). By adjusting the digital down-conversion frequency, DDS can also be applied in the AFC procedure. It should be remembered, however, that the analog channel selection filter defines the adjustment margin. Thus, it might be useful to define its bandwidth wider than the final desired channel BW and perform the final selection digitally. If, on

the other hand, a wide IF filter with high dynamic range ADC is used, the DDS can make the actual frequency channel selection. This may not be as useful with wideband systems as with narrowband ones and would clearly be a base station oriented solution.

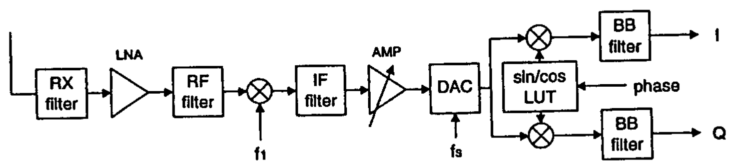


Figure 10.24 1-stage super-heterodyne receiver with digital I/Q splitting.

10.5.2.2 *Direct Down-Conversion*

The direct down-conversion receiver shown in Figure 10.25 mixes the receiver RF signal straight into I/Q baseband components for which the channel filtering is performed using simple, cheap low-pass filters. The benefit of this architecture is the reduced complexity and performance, especially in terms of the SAW filter (high insertion loss), and may in the end lead to a single chip transceiver including also the transmitter section.

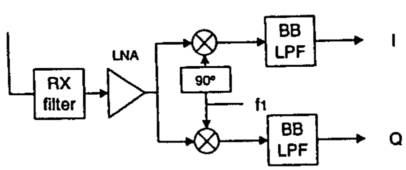


Figure 10.25 Direct-conversion receiver with baseband channel filtering.

There are, however, a few problems related to the direct conversion receiver technique. One aspect is the DC component, which may appear because of self-mixing the local frequency and/or nearby strong users. A CDMA signal inherently is resistant against narrow band interferers, but high level DC offset could saturate the input ADCs and destroy the reception.

The key component here is the quadrature local oscillator, which should provide good phase and gain matching in order to provide pure I/Q-branch sine waves with

equal amplitude and 90-degree phase shift. Still, the isolation from the mixer input must be high enough to avoid the above-mentioned self-mixing. The mixer itself should possess good linearity to suppress intermodulation products caused by adjacent/alternate channel transmission.

Finally, it should be noted that the direct conversion receiver can be fit into a single chip by using, for example, the silicon bipolar process (in the future, a CMOS process could also be applied). The low pass filters are excluded from this and are put into the baseband side.

10.5.2.3 *Harmonic IF Sampling*

So far it has been assumed that the receiver I/Q baseband signal is converted into the digital domain using a sampling frequency twice as high as the highest spectral component in the input. The Nyquist criteria states, however, that it is enough to use a sampling frequency twice the desired signal bandwidth, which suggests that it is fully possible to utilize the actual signal alias term. In the frequency domain, the desired band is translated to another phase according to the selected sampling rate. Hence, the sub-sampling can also be considered as a down-conversion process (see Figure 10.26). Before feeding the signal to an A/D converter, bandpass channel selection filtering must be done for the desired signal to prevent other channels to cause their aliases within the desired frequency band.

The ADC can be used as such if the sampling rate suits the signal central frequency. If not, a separate track/hold function with a low pass filter should be applied to relax the ADC requirements. Maximum input signal frequency to the sampling device depends on the track and hold device bandwidth. In practice, several hundreds of megahertz can already be handled.

From the CDMA point of view, sub-sampling receiver does not give any benefits over traditional narrowband systems. From a implementation point of view, however, only one ADC with fast T/H instead of two ADCs is required. Complexity on the digital side is increased due to I/Q separation that must be done at a high sampling rate.

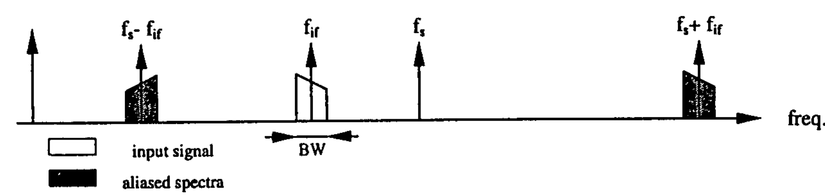


Figure 10.26 Signal spectrum after performing subsampling.

10.5.3 Transmitter Architectures

This section presents the basic RF transmitter architectures suitable also for wideband CDMA transceiver implementations. Critical aspects are highlighted and, in conjunction with each approach, advantages and disadvantages are revised.

The baseband signal to be transmitted is converted to the final carrier either directly or through some intermediate frequency. The latter approach has been more common in the past, while present RF technology is capable of following the former approach.

In direct sequence spreading systems, phase and/or amplitude shift keying chip-level modulation is generally utilized. A single sideband signal with suppressed carrier is generated using an I/Q-mixer with 90-degree phase difference in the frequency between in-phase and quadrature local inputs.

10.5.3.1 *Traditional IF-Based Up-Conversion*

Transmitted signal conversion to intermediate frequency before the final carrier provides a few advantages. First, different stages are easier from a component matching point of view. An I/Q modulator, for example, operating with low frequency can offer better phase and amplitude balance. One important aspect is that channel filtering can also be done using IF passband filters and consequently relax the processing power requirements of the digital parts.

The main sources of noise come from the nonideal local frequencies. This may be problematic if there are several IF stages because each local oscillator (LO) adds to the total transmitter noise contribution.

CDMA systems demand tough requirements for the dynamic range, accuracy and speed of the transmission power control. In practice, the amplification/attenuation is done in many phases to fulfill these requirements and an IF-based transmitter has an intrinsic feature to support gain control in several stages.

A functional block diagram of a 1-stage IF transmitter is shown in Figure 10.27. The I/Q -signal is up-converted to a predefined first IF frequency that is likely to be constant. This is followed by passband filtering and an IF amplifier. Finally, the signal is translated to carrier frequency, amplified by the PA, and input to the transmission RF filter. A duplex filter is required in case of simultaneous transmission and reception. The highest up-conversion frequency is traditionally made variable and is used in the transmission channel selection.

DDS techniques can be utilized with an IF-based transmitter. In the simplest form, DDS replaces the I/Q mixer or even the analog RF synthesizer. The disadvantage of this approach is that the generated frequencies are not as clean as the analog counterparts. The advantage is the flexibility in terms of frequency and settling speed.

By moving the digital-analog border closer to the antenna, the arrangement shown in Figure 10.28 is achieved. Here the I/Q -modulation is done digitally before converting to analog format. Good amplitude and phase balance can be obtained, but the disadvantages noted before still hold. Digital-to-analog converter (DAC) also causes

distortion due to the S/H-function, which must be taken into account when designing the reconstruction filter.

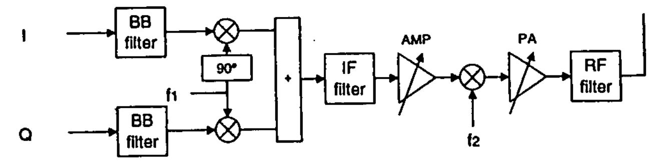


Figure 10.27 Typical IF transmitter architecture.

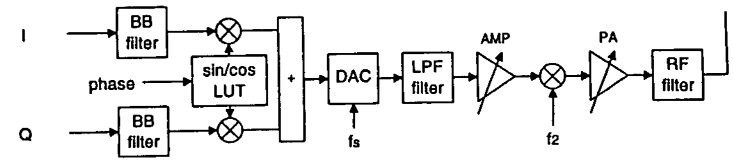


Figure 10.28 IF based transmitter with digital I/Q -modulation.

10.5.3.2 *Direct Up-Conversion*

The main disadvantage of a several-stage transmitter is the number of different components required, which increases the cost of a portable terminal. In a direct conversion-type transmitter (Figure 10.30), the signal is transferred directly from baseband to RF band. There is a need only for one synthesizer whose output is taken to the I/Q -modulator. Care must be taken to provide good enough 90-degree phase offset and gain match between the quadrature branches in order not to create too low an unwanted sideband suppression.

Channel filtering in the case of a direct up-conversion transmitter must be performed on the baseband before I/Q mixing. Digital filtering with linear phase, and hence constant group delay, is a good option for the purpose. The analog low pass filters following the DAC could also be used in the process.

Power control requirements are slightly more difficult to fulfill in case of direct up-conversion. Some alleviation can, however, be provided by implementing higher

dynamic range digital parts and especially the D/A-converter. From the implementation point of view, higher dynamic range digital parts, and especially the D/A-converter, are needed.

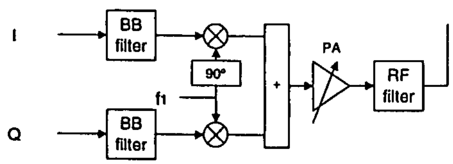


Figure 10.29 Direct conversion transmitter.

10.6 SOFTWARE CONFIGURABLE RADIO

The most flexible implementation for a communication terminal would incorporate a software configuration. Such an approach could provide an access to several different wireless systems with a single hardware unit simply by downloading a new configuration internally or from a network. The idea of a software radio has originally been connected to military applications. Such a device, however, would provide flexible roaming between today's multitude of commercial cellular systems. Good articles about software and multimode radios can be found in [32–34].

Figure 10.30 illustrates the concept of the software radio. There are a few basic features that relate to the software radio. The RF section must support transmission and reception of differently modulated signals within several frequency bands. This calls for a linear power amplifier operating over a wide frequency range with acceptable efficiency. Similarly, antennas must provide low loss and uniform gain across the same range. After down-conversion to IF or baseband initial channel, selection prior to the A/D conversion must be performed. This is necessary in order to restrict the signal dynamic range to within acceptable limits. If a large dynamic range A/D converter and a wide enough filter are used, the actual channel filtering can be made digitally. This may not, however, be enough if the nearby channel powers are radically larger than the desired channel. One solution here is to introduce a filter bank and select one bandwidth according to the operations mode.

The baseband section of a software radio consists of programmable hardware and a powerful DSP processor(s). In the former case, the configuration can be done by employing field programmable gate array (FPGA) technology and well defined hardware description languages like VHDL (Very High Speed Integrated Circuit Hardware Description Language) or VERILOG. Depending on the access method, different baseband functions need to be supported. The main concern here is in the

channel equalization and forward error-correction coding/decoding. The baseband architecture must be designed in such a way that both burst- (TDMA) and symbol-type (DS-CDMA) equalization approaches can be supported in a straightforward manner. When it comes to channel encoding/decoding there are numerous methods available, each providing certain coding rates and service qualities. Software implementation is fine with low effective bit rate systems, but high rate systems call for hardware accelerators.

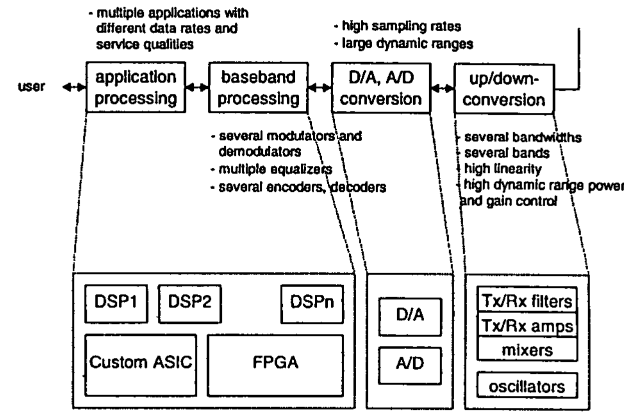


Figure 10.30 Mobile station software radio top-level architecture.

The application side of a mobile terminal has so far been mostly speech oriented and several complex source codecs have been implemented into the handsets. New applications related mainly to video coding and other data processing are emerging, and hence, the requirements for the software radio application support become greater.

An important issue with the software radio is related to flexible usage of the available processing resources. The system clock rate cannot be increased easily in a mobile terminal, and hence, one of the key solutions is to utilize configurable silicon. Higher processing capabilities are obtained by creating and activating extra processing units on a need basis. This could be interpreted as a "breathing" parallel processor that expands as needed. Power consumption is occasionally increased and must be taken into account while designing the device power management.

10.7 TYPICAL WIDEBAND CDMA MOBILE TERMINAL ARCHITECTURE

This section presents an architecture for a wideband CDMA mobile terminal integrating the different functions presented earlier. For the RF section, a traditional down-conversion receiver architecture with two IFs was selected, and on the transmitter side,

direct up-conversion will be employed. The baseband side consists of all the bit- and chip-level processing primitives needed to provide all bearer service quality classes, band spreading/despreading, coherent multipath diversity detection, and some extra functions for link maintenance.

Figure 10.31 shows the mobile transmitter section. On the transmitter side, the information-bearing user data is obtained from outside the radio modem part. The stream is coded and interleaved according to selected QoS. Two basic encoding schemes, namely convolutional and Reed-Solomon, are available. The rate of the user data is matched to the air interface by choosing some of the available spreading ratios and unequal symbol repetition. Frame control header information about the frame data and unequal symbol repetition. Frame control header information about the frame data rate and coding scheme is encoded and interleaved to the quadrature branch. In addition, the power control and the required amount of reference symbols are placed into the transmitted stream. Before D/A conversion, the physical in-phase and quadrature branches are spread with two PN codes and passed through a pulse shaping filter to obtain the specified transmission spectrum.

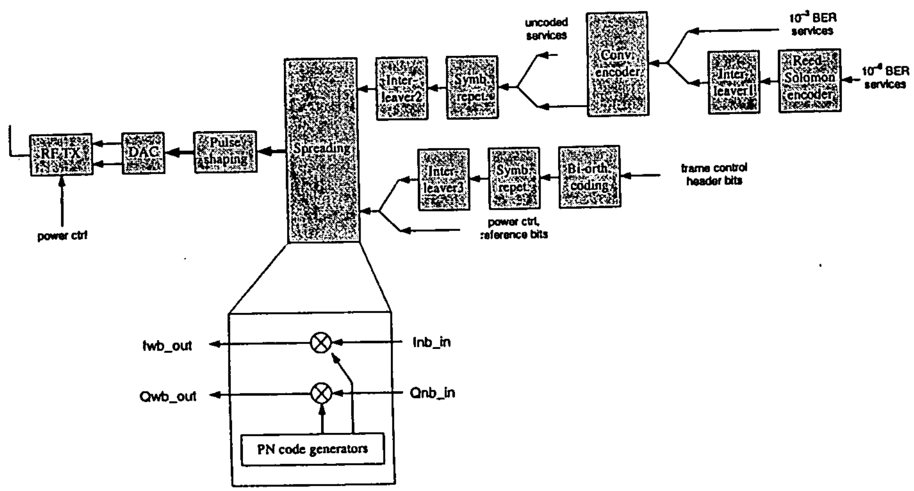


Figure 10.31 Mobile transmitter section.

The receiver side down-converts the incoming signal to the baseband before A/D conversion. Figure 10.32 depicts the mobile terminal receiver section. Channel filtering and AGC have been implemented on the analog RF side. In addition, a loop back from the baseband channel estimator provides an error signal to adjust the reference oscillator for clock synchronization purposes. The I/Q signal is converted to

the digital domain using a 6-bit, four times oversampling ADC. Digital filtering following the ADC is matched to the transmitter filter, thus improving the SNR. The digital filter can also be used to compensate for nonlinearities caused by the analog side filters. Wideband power measurement is performed for the post-ADC signal and is used in the receiver AGC loop. The multipath channel estimator utilizes a full matched filter to find multipath components quickly and to perform accurate delay estimation without any DLL. RAKE fingers consisting of only two physical correlators perform the signal despreading for each multipath component and parallel 20 ksymbol/s channel. A maximal ratio combiner estimates complex channel taps from the despread pilot channel, rotates/weights the RAKE output signals, and sums them together. In addition, the combiner outputs signals for frequency error, narrowband signal power, and SIR. These are used in AGC, AFC, and PC-loop, as described earlier. The selector picks up the code channels (I/Q branches) owned by the specific user and routes to de-interleavers/channel decoders.

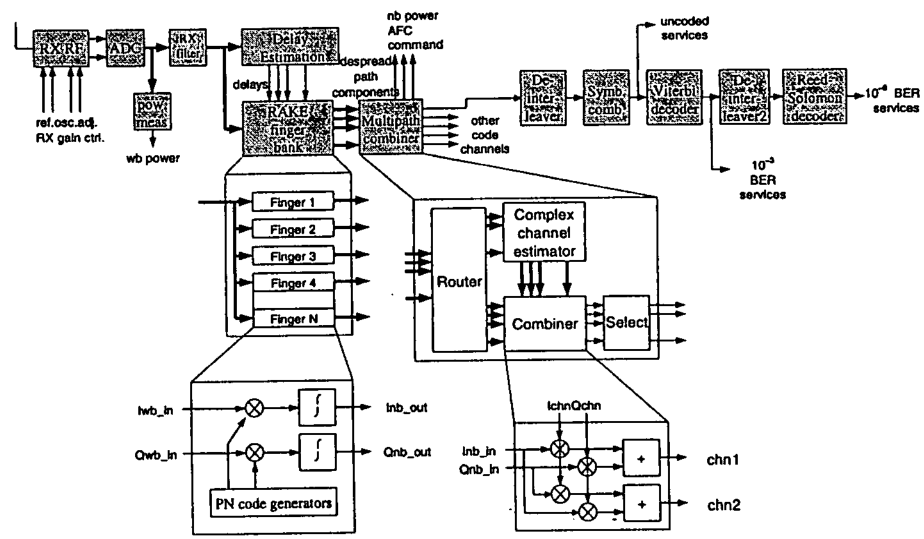


Figure 10.32 Mobile terminal receiver baseband section.

10.8 MULTIMODE TERMINALS

Since at the beginning of wideband CDMA system deployment there will still be several second generation systems in operation, dual-mode or multimode terminals can be used to provide seamless service for the users. However, from the transceiver point

of view, this may be cumbersome because the same physical terminal must provide technical solutions to support potentially dissimilar radio technologies.

The main second generation cellular systems are GSM, IS-136, IS-95, and PDC. Furthermore, PHS and DECT are second generation low-tier wireless systems. Third generation wideband CDMA systems will be implemented together with some of these second generation systems. The most likely combinations are GSM/WCDMA, PDC/WCDMA, and IS-95/cdma2000.

Different systems place different requirements on different parts of the transceiver. On the other hand, the right selection of transceiver architecture can ease the implementation of dual mode terminals. In the RF section, the duplexer is impacted by the fact that second and third generation systems have different frequency bands. RF and IF filter bandwidths are different due to different bandwidths of the systems. Intermodulation and phase noise requirements are also most likely different. Synthesizer settling times vary, as well as A/D and D/A converter requirements. Since different systems have different symbol rates and other clocking requirements, possibly two reference oscillators are required. This could be avoided by proper selection of system parameters for the new system. In the baseband, receiver algorithms are somewhat different. For example, a TDMA-based system requires an equalizer and a CDMA-based RAKE receiver.

Since there are so many different dual-mode combinations, it is impossible to analyze each of them in detail. Instead, several commonalties can be observed, and implementation of dual-mode terminals can be analyzed in general. The main difference of the above-listed systems is the spreading used in CDMA systems and the continuous transmission compared to nonspread, slotted signal of TDMA-based systems. Therefore, the text below assumes separation between nonspread and spread systems, between slotted and continuous systems, and according to a duplex scheme.

10.8.1 Nonspread Versus Spread Systems

The difference in the implementation of nonspread and spread systems is certainly not straightforward. The major parameter is the channel bandwidth, and if these are similar for both the systems, the actual transceivers seem fairly unified. The major differences are on the baseband transmitter and receiver front-ends. In the former case, the spread system faces a multilevel signal due to parallel code channels, which increaseases transmission filter and D/A converter requirements. On the receiver side, the nonspread equalizer (DFE or MLSE) is also replaced with a RAKE receiver, or even a multiuser detection unit.

If the bandwidths are dissimilar, like with wideband CDMA and GSM, the situation is different. First, the IF/RF section will require new filters on both TX/RX sides. For a dual-mode terminal, there should be several filters from which the correct one is selected. Another choice is to implement an adjustable filter, which has the disadvantage that it cannot always be adjusted the most optimal way.

Channel raster is directly related to the bandwidth used. If the specified rasters are not factors of each other and analog synthesizers are used, there is a problem with

the reference frequency. In order to provide simple and fast PLLs, the reference frequency should equal the channel raster. For a multimode terminal this would mean a compromise between complexity (possible several references) and speed (lowest common factor frequency).

On the baseband side the processing requirements between spread and nonspread systems can be considered fairly similar. A wideband signal calls for a higher sampling rate, while at the same time it can be presented with less bits/sample. In favor of nonspread systems, signal spreading/despreading adds some extra complexity to the terminal. For long channel delay spreads, a RAKE receiver tends to be simpler than an optimal MLSE equalizer.

10.8.2 Slotted Versus Continuous Systems

Continuous systems better suit ASIC implementation because most of the functions operate constantly. Selecting a software approach for a continuous system would cause high control overhead per received symbol because each one would be processed independently by generating an interrupt for a DSP. This could be overcome by buffering some number of symbols prior to transmission and reception. The problem here, however, is the increased data path latency.

Slotted systems on the other hand suit software processing fairly well. This is because symbol detection may start only when the whole/partial burst has been received. This is especially the case when the channel estimation is performed from the burst midamble sequence. Overhead due to control is minimized, and the number of instructions per processed symbol becomes feasible.

From a slotted systems' RF point of view, the power amplifier sees discontinuous transmission with high peak power. With continuous systems, the PA is active constantly, and, if the average transmission power is assumed the same, the peak power is reduced respectively. Continuous systems may, however, act as discontinuous systems if DTX mode is applied. This is common especially with speech services.

10.8.3 FDD Versus TDD Systems

Continuous systems generally use frequency duplex division for providing up and downlink channels. From an implementation point of view, this requires synthesizers on both the transmitter and receiver side due to simultaneous operation. In addition, a duplex filter must be applied to prevent TX signal leaking to RX side.

In slotted systems, another possible method is to utilize time division duplex separation. The idea is to transmit up and downlink channels with the same frequency but at different times. The main benefit here is that the terminal is not transmitting/receiving simultaneously, and hence only one synthesizer and no expensive duplex filter is needed.

On the baseband processing side there are few physical resources that could be shared between the transmitter and receiver. DS-CDMA code generators are such an example (polynomials, and states reloaded upon TX/RX switching). In addition, the

multipath channel estimation could be made easier because the channel is reciprocal (i.e., the channel estimate obtained for one direction is valid within coherence time for the other direction).

REFERENCES

[1] Neuvo, Y., "Wireless Will Catch Up Wireline and More", *Keynote speech in PIMRC'97*, Helsinki, Finland, September 1997.

[2] Bernel, D. A., and T. C. Hofner, "Digital Comms. Require Dynamic Performance," *Electronic Engineering*, September 1997, pp 67–71.

[3] Gaudenzi, R. D., "The Influence of Signal Quantization on the Performance of Digital Receivers for CDMA Radio Networks," *ETT*, Vol. 8, No. 1, January - February 1997, pp. 89–97.

[4] Do, G. L., "Efficient Filter Design for IS-95 CDMA Systems," *IEEE Trans. on Consumer Electronics*, Vol. 42, No. 4, 1996, pp.1011–1019.

[5] Dixon, R. C., *Spread Spectrum Systems with Commercial Applications*, New York: John Wiley & Sons Inc., 1994.

[6] Pickholz, R. L., D. L. Schilling, and L. B. Milstein, "Theory of Spread-Spectrum Communications – A Tutorial," *IEEE Trans. on Comm.*, Vol. 30, No. 5, May 1982, pp. 855–884.

[7] Polydoros, A., and C. L. Weber, "A Unified Approach to Serial Search Spread-Spectrum Code Acquisition – Part 2: General Theory," *IEEE Trans. on Comm.*, Vol. 32, No. 5, May 1984, pp. 550–560.

[8] Bensley, S. E., and B. Aazhang, "Subspace-Based Channel Estimation for Code Division Multiple Access Communication Systems," *IEEE Trans. on Comm.*, Dec. 1994, pp. 1009–1020.

[9] Kreyszig, E., *Advanced Engineering Mathematics*, 7th ed., New York: John Wiley & Sons Inc., 1993.

[10] Lupas, R., and S. Verdú, "Linear Multiuser Detectors for Synchronous Code-Division Multiple Access Channels," *IEEE Trans. on Information Theory*, Vol. 35, No. 1, January 1989, pp.123–136.

[11 Paris,] B. P., "Finite Precision Decorrelating Receivers for Multiuser CDMA Communication Systems," *IEEE Trans. on Comm.*, Vol. 44, No. 4, April 1996, pp. 496–506.

[12] Xie, Z., R. T. Short, and C. K. Rushforth, "A Family of Suboptimum Detectors for Coherent Multiuser Communications," *IEEE Journal on Selected Areas in Communications*, Vol. 8, No. 4, May 1990, pp. 683–690.

[13] Varanasi, M. K., and B. Aazhang, "Multistage Detection in Asynchronous Code Division Multiple-Access Communications," *IEEE Trans. on Comm.*, Vol. 38, No. 4, April 1990, pp. 509–519.

[14] Holzman, J. M., "DS/CDMA Successive Interference Cancellation," *Proceedings of ISSSTA'94*, Oulu, Finland, July 1994, pp. 69–78.

[15] Patel, P., and J. Holtzman, "Analysis of a Simple Succesive Interference Cancellation Scheme in a DS/CDMA System," *IEEE Journal on Selected Areas in Communications*, Vol. 12, No. 5, June 1994, pp. 796–807.

[16] Giallorenzi, T. R., and S. G. Wilson, "Decision Feedback Multiuser Receivers for Asynchronous CDMA Systems," *Proceedings of Globecom'93*, Houston, Texas, USA, November 1993, pp. 1677–1682.

[17] Duel-Hallen, A., "A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple Access Channels, " *IEEE Comm*, Vol. 43, No. 2/3/4, February/March/April 1995, pp. 421–434.

[18] Lippmann, R. P., "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, April 1987, pp. 4–19.

[19] Paris, B. P., B. Aazhang, and G. Orsak, "Neural Networks for Multi-User Detection in CDMA Communications," *IEEE Trans. Comm*, Vol. 40, July 1992, pp. 1212–1222.

[20] Kechiotis, G., and E. S. Manolakis, "Hopfield Neural Network Implementation of the Optical CDMA Multiuser Detector," *IEEE Trans. Neural Networks*, January 1996, pp. 131–141.

[21] Kechiotis, G., and E. S. Manolakis, "A Hybrid Digital Signal Processing-Neural Network CDMA Multiuser Detection Scheme," *IEEE Trans. on Circuits and Systems-II*, Vol. 43, No. 2, February 1996, pp. 96–104.

[22] Proakis, J. G., *Digital Communications*, 3rd ed., New York: McGraw-Hill, 1995.

[23] Dawid, H., G. Fettweis, and H. Meyr, "A CMOS IC for Gb/s Viterbi Decoding: System Design and VLSI Implementation," *IEEE Trans. on VLSI Systems*, Vol. 4, No. 1, March 1996, pp. 17–31.

[24] Anderson, J. B., and E. Offer, "Reduced-State Sequence Detection with Convolutional Codes," *IEEE Trans. on Information Theory*, Vol. 40, May 1994, pp. 944–955.

[25] Forney, G. D., "The Viterbi Algorithm," *Proc. of IEEE*, Vol. 61, March 1973, pp. 268–278.

[26] Viterbi, A. J., and J. K. Omura, *Principle of Digital Communication and Coding*, New York: McGraw-Hill, 1979.

[27] Blahut, R. E., *Theory and Practice of Error Control Codes*, Reading, MA: Addison-Wesley Publishing Company, 1993.

[28] Oppenheim, A. V., and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall International Inc., 1989.

[29] Sundström, L., "RF Amplifier Linearisation Unsing Digital Predistortion," *Thesis for the degree of Teknisk Licentiat*, Lund University, Sweden, June 1993, pp. 8-17.

[30] Johansson, M., and M. Faulkner, "Linarization of wideband RF power amplifiers," *Proceedings of Nordic Radio Symbosium*, April 24–27, 1995.

[31] Cavers, J. K., "A Linearising Predistorter with Fast Adaptation," *Proceedings of 40th IEEE Vehicular Technology Conference*, May 1990, pp.41–47.

[32] Special issue on software radios, *IEEE Communications Magazine*, Vol. 33, No. 5, May 1995.

[33] Kuisma, E., "Technology Options for Multi-Mode Terminals," *Mobile Communications International*, February 1997, pp. 71–74.
[34] Edmond, P., "Towards a multi-mode mobile telephone design," *Electronics Engineering*, October 1996, pp. 67–70.

Chapter 11

NETWORK PLANNING

11.1 INTRODUCTION

When using a cellular phone, you have most likely experienced several undesired effects such as dropped calls when you move from one location to another, busy network when attempting to make a call, or just poor speech quality. The grade of service wireless subscribers experience is dependent on the quality of the radio network planning. Thus, network planning is one of the key competence areas for any operator in order to satisfy customers and thus make a profit.

Network planning covers two major areas: radio network planning and network dimensioning. Radio network planning includes the calculation of link budget, capacities, and thus the required number of cell sites. Furthermore, radio network planning includes detailed coverage and parameter planning for individual sites. In network dimensioning, the required number of channel elements in the base station, the capacities of transmission lines, the number of base station controllers, switches, and other network elements are calculated. This chapter presents the main concepts of third generation network planning; for details of IS-95 network planning refer to [1].

A large number of different bit rates and the diversification of services complicate the network planning process for third generation systems, as compared to the second generation systems. It is more difficult to predict traffic and usage patterns of different services. Which unit to use to measure traffic density is our first topic in this chapter. Next, we introduce performance measures, such as outage and blocking, which are used to measure the network quality. High bit rates services can be offered either with uniform coverage over a cell or with smaller data rates at the cell edge. Furthermore, either continuous coverage over a wide area or only hotspot coverage could be provided. A careful assessment is required to understand what is the impact of different deployment strategies on the demand of services.

The radio network planning process can be divided in three phases: preparation, estimation of cell count, and network optimization. Each of these is discussed in depth.

In the network optimization phase, the following aspects are discussed: detailed characterization of the radio environment for individual cell planning purposes, CDMA control channel power planning, pilot pollution, planning of soft handover parameters, interfrequency handover, iterative network coverage analysis, and radio network testing. The impact on the network planning of microcell and indoor planning are discussed, as are sectorization and beamforming.

The main emphasis in the section on network dimensioning is on the calculation of the number of channel elements in the base station.

The co-existence of wideband CDMA air interface with already deployed air interfaces will be an important factor in the beginning of wideband CDMA system deployments. The last sections discuss intermodulation interference, guardband and zones, and transition considerations.

11.2 TRAFFIC INTENSITY

The starting point for network planning is the assessment of individual user traffic and traffic intensity (offered traffic) in a given area. To quantify the traffic intensity, different units can be used. Most often, traffic intensity is measured in *Erlang*, where 1 Erlang is equivalent to one circuit in use for 1 hour (3600 seconds).

$$\text{Traffic intensity in Erlang} = \frac{\text{Number of calls per hour} \times \text{average call holding time (seconds)}}{3600} \tag{11.1}$$

For example, if a user makes one call per hour and the average call duration of calls is 120 seconds, it produces 120/3600 = 33 mErlang of traffic. Given the user density in a geographical area, the traffic density can be calculated and is expressed in Erlang per square-kilometer (Erl/km^2).

Since third generation systems will have a very large variety of services, a single traffic measure might not be suitable to all cases. The *equivalent telephony Erlang* (ETE) defines traffic with relation to basic telephone calls [2]. However, this definition is dependent on the transmission rate of the basic telephone service. For data services, traffic measured in Mbps/km^2 will better characterize the traffic density.

11.3 PERFORMANCE MEASURES

The performance measures for the grade of service are *area coverage probability* and *blocking*. The area coverage probability is related to the quality of the radio planning and the radio network capacity. Blocking is related to the available hardware, for example, to the base station channel cards. The term *soft blocking* is used when a user is blocked because the cell runs out of capacity (i.e., a user is denied access to the network in order to ensure the quality of already admitted users). First, the network is designed for the desired area coverage probability. Within this area coverage probability, then, certain blocking requirements, usually 2%, need to be fulfilled.

Area coverage probability can also be defined by *outage*. Outage is the probability that the radio network cannot fulfill the specified quality of service target. Several definitions exist for the measure of outage. In general, it is defined as a drop in the required quality below a prespecified target value. However, it needs to be defined for how long the quality must be below the target value to create an outage. Furthermore, quality criteria for outage must be defined. We can separate between a *signal outage*, when the signal-to-noise ratio (SNR) drops below the target value, and *interference outage*, when the signal-to-interference ratio (SIR) drops below the target value. The outage is impacted by several factors, such as shadowing, pathloss model, outside cell interference, handover type, and power control dynamics, and is usually evaluated by system level simulations.

If the system is *coverage limited*, outage can then be defined as the probability that pathloss and shadowing exceed the difference between the maximum transmitted power and the required received signal level. Since shadowing is log-normally defined, very low outage probability leads to very high power margin and thus small cells and an expensive network. This is illustrated in Figure 11.1, which shows a plot of a log-normal distribution with standard deviation of 8 dB. The area of the tail of the distribution gives the outage probability. In order to achieve 10% outage probability (i.e., 90% *availability* or coverage probability), a 10.3-dB shadowing margin is required. For 5% outage target, a 13.2-dB margin is required, and for 1% outage target, an 18.6-dB margin. Thus, we can see that increasing the outage target leads to very high margins and thus reduced range. We also plot shadowing with the standard deviation of 10 dB. The increased shadowing requires higher margins since it results in higher probability of large shadowing values.

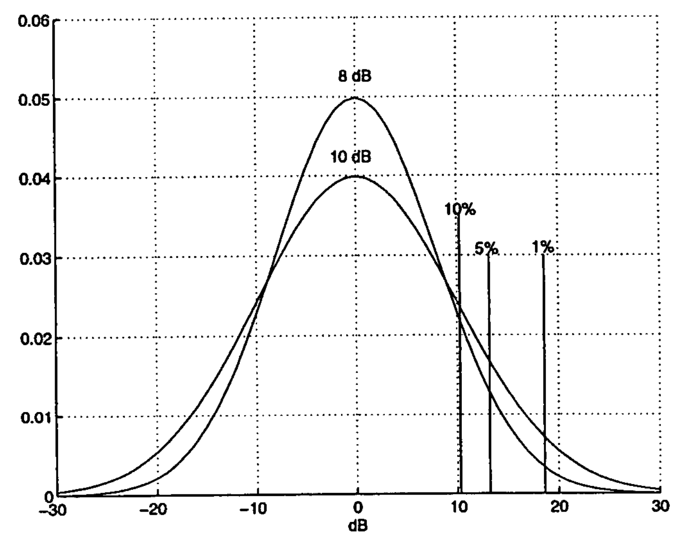


Figure 11.1 Log-normal distribution and shadowing margin.

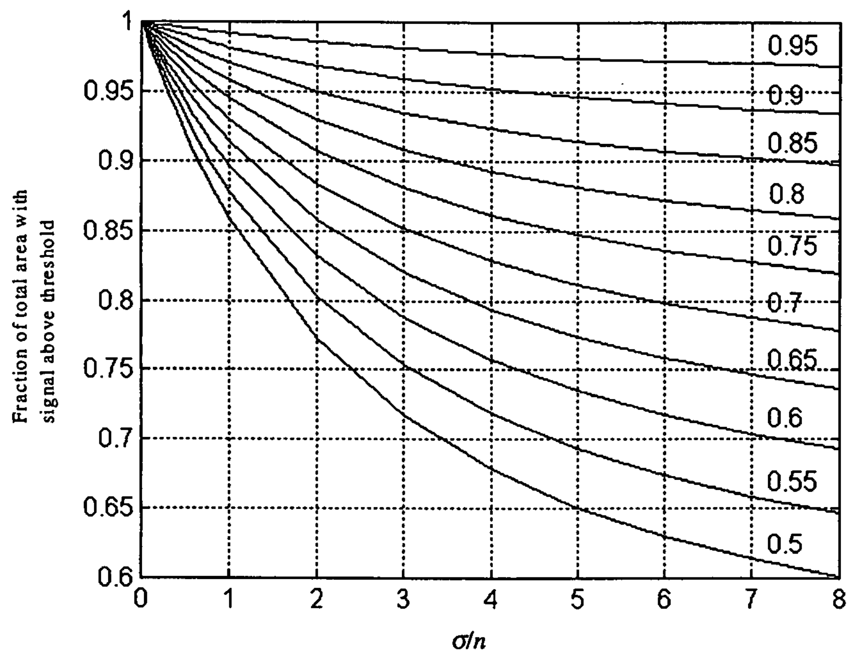


Figure 11.2 Curves relating fraction of total area with signal above threshold as a fuction of probability of signal above threshold on the cell boundary.

The coverage probability can be defined as a cell edge probability or a cell area probability. The cell edge probability refers to a probability that the mobile would receive a signal above a threshold at the cell edge, for example, RF signal strength. The threshold value chosen need not be the receiver noise threshold but may be any value that provides an acceptable signal under Rayleigh fading conditions [3]. The cell area probability means the percentage of locations within a circle of radius R in which the received signal from a radiating base station antenna exceeds a particular threshold value. The relation between area and cell edge probability has been derived in [3] (pp. 126-127) and is illustrated in Figure 11.2, where n is the pathloss exponent and σ is the standard deviation of the shadowing given in decibels. In a macrocell environment, a typical value for the pathloss exponent is 3.6 and for standard deviation is 8 dB. Thus, σ/n is 2.22, and a boundary coverage probability of 85% would give 95% area coverage.

Blocking is defined as the probability of a blocked call on a first attempt. A blocked call occurs due to a lack of system resources (i.e., if all network resources are busy, a user will be denied access to the system). The reasons for blocking can be shortage of channel elements in the base station or shortage of fixed network resources

such as transmission lines, BSC, or switching capacity. Thus, each network element needs to be dimensioned appropriately for the traffic demands. Usually, a blocking probability of 2% is used when dimensioning cellular systems.

Let N be the total number of channels and T be the offered traffic in Erlang. Then, the probability of all channels being busy is obtained from a Poisson distribution:

$$P(N;T) = \frac{T^N e^{-T}}{N!} \tag{11.2}$$

where $P(N;T)$ is the blocking rate. Thus, if the traffic density (offered traffic) and blocking probability are known, the required number of network resources (e.g., channel elements in base station or capacity of the transmission lines) can be calculated using the Erlang tables. The Poisson distribution has the effect of achieving better trunking efficiency as the number of channels is increased. Trunking efficiency, also called channel utilization efficiency, is defined as

$$\text{Efficiency}(\%) = \frac{\text{Traffic in Erlang}}{\text{Number of channels}} \times 100\,\% \tag{11.3}$$

Given the blocking probability, (11.2) can be used to calculate the traffic in Erlang for different numbers of channels. Thereafter, (11.3) gives the trunking efficiency. In Figure 11.3, we have plotted the trunking efficiency for a blocking probability of 2%. We can clearly see that the efficiency increases as the number of channels increases.

11.4 RADIO NETWORK PLANNING PROCEDURE

The prerequisite for good radio network planning is the know-how of the radio environment. Rough network planning and deployment plans can be made based on general radio channel models. However, since the radio environment is highly variable, even within the area of one cell, detailed measurements and optimization need to be performed for each individual cell.

The radio network planning process can be divided into three phases:

- Preparation;
- Estimation of cell count;
- Detailed network planning.

11.4.1 Preparation Phase

In the preparation phase, coverage and capacity objectives are established, network planning strategy defined, and initial design and operating parameters determined.

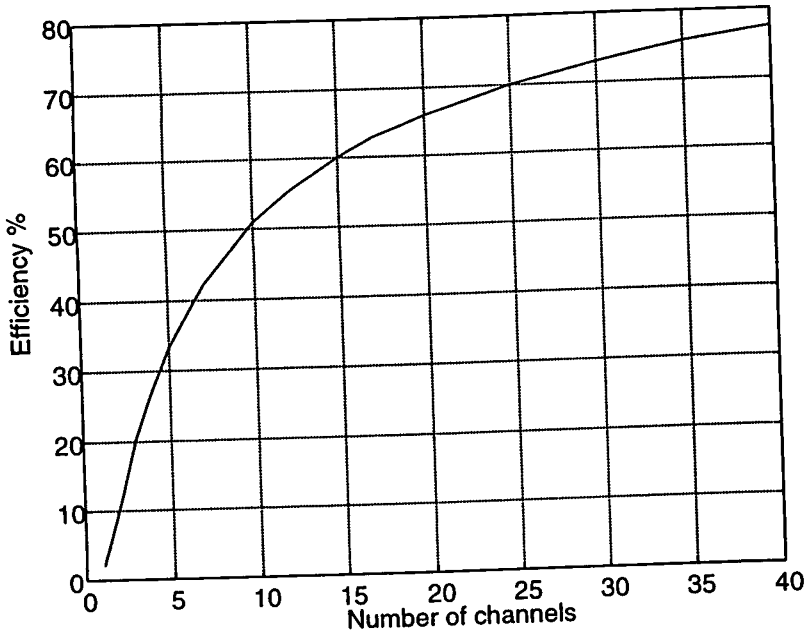


Figure 11.3 Trunking efficiency.

11.4.1.1 Coverage and Capacity Objectives

Coverage and capacity objectives are a trade-off between desired quality and overall network cost. A smaller signal outage probability means smaller cells and thus higher overall network costs; and smaller interference outage probability means smaller capacity and thus also higher cost. A typical outage probability target is 5% to 10%, corresponding to 90% to 95% availability/coverage probability. The coverage probability could be different for different services as discussed below.

11.4.1.2 Network Planning Strategy

The network planning strategy includes issues like microcell deployment, provision for indoor and high bit rate coverage, and migration from second generation systems. Several factors need to be considered for the most feasible network planning approach. These include cost of fixed line transmission, how easily cell sites can be acquired, and at what cost cell sites can be acquired. Furthermore, environmental issues such as cell tower appearance will impact where a base station can be deployed. Traffic distribution will of course impact the deployment strategy.

One deployment strategy could be to use macrocells for outdoor coverage and picocells for indoor coverage in office buildings. In addition, macrocells would be used to fill the gaps in the indoor coverage. This is because extensive indoor coverage is most likely required in any case. Therefore, it might be wiser to build additional capacity by increasing the number of indoor cells rather than trying to provide indoor coverage from outdoor cells and, thereby, be forced to introduce microcells earlier due to capacity restrictions. If coverage in indoors is provided by outdoor base stations, the building penetration margin, typically 10 to 20 dB, needs to be taken into account in the link budget calculations.

Another approach is to use microcells extensively from the beginning and to provide indoor coverage from them. This might be feasible in dense urban areas.

High bit rates can be provided either uniformly over the cell area or the data rate at the cell border could be smaller than when close to the base station to allow a larger cell range. This depends on the nature of the high bit rate services. For services which utilize available bit rates (i.e., services that do not require any quality of service guarantees), nonuniform coverage might be acceptable, but for applications that require maximum bit rate such as video transmission, uniform coverage is required.

If an operator has a deployed second generation network, migration aspects need to be considered in the network planning strategy. These include reuse of existing cell sites, handovers between the new and old systems, and co-existence requirements (see Section 11.9).

11.4.2 Estimation of Cell Count

The number of wireless users in a given area is obtained by multiplying the population by the penetration. The number of users and offered traffic per user determine the overall offered traffic. When cell capacity and range are known, a rough number of cells can be determined. Figure 11.4 illustrates the procedure for the estimation of cell count.

11.4.2.1 Offered Traffic

In the calculation of the number of users that need to be served, at least the following factors should be accounted for:

- Population living in given area;
- Population working in given area;
- Vehicle traffic;
- Special events, and use of recreational areas.

All the above aspects contribute their share to the total number of users that might use the cellular network. When the total number of users is known, we need to estimate what percentage of the total number of users has a wireless terminal (i.e., what is the penetration?).

The offered traffic per user depends on what services are used and how often they are used. The prediction of service usage is especially difficult for third generation

for which many new services will be introduced. Different services produce different bit rates. For data services, we need to estimate the average requirement of megabit per second per user. The average calling time for circuit switched users depends on the service (e.g., speech, video), customer types (business, residential), and tariffing. Typical values for the average calling time for speech service range from 120 to 180 seconds. For other services, the average calling time is more difficult to define since there is less experience. Since the average calling time has a major impact on the generated network traffic, it is very important to be able to predict it with reasonable accuracy.

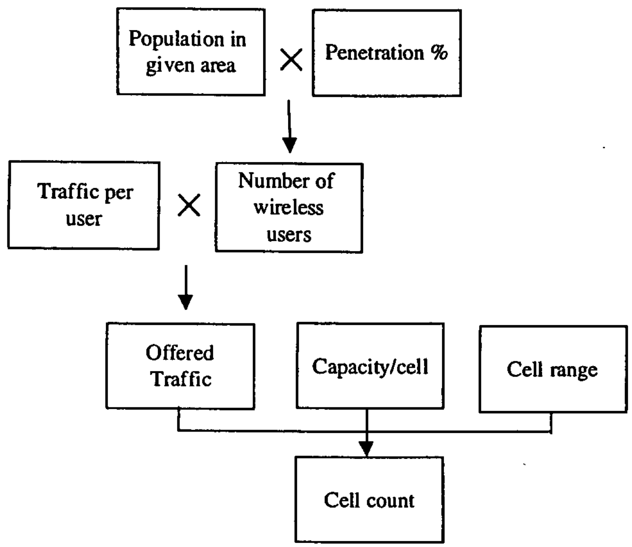


Figure 11.4 Estimation of cell count.

Since the demand for services also varies according to the time of the day, the network capacity must be dimensioned according to busy hour traffic (i.e., when the traffic demand is highest). Typically, busy hour is during business hours. However, new user types (e.g., young people) and a more widespread use of entertainment services through wireless connections might change this.

Example: The total user density is 10,000 user/km^2, 50% of users are using speech and 10% data. Thus, the number of users requiring speech services is 5000 user/km^2 and data services 1000 user/km^2. A speech user generates 30 mErlang traffic and a data user 100 mErlang. Thus, the total offered traffic for speech is 150 Erlang and for data 100 Erlang. If a data call requires on average 100 Kbps, the total offered data traffic is 10 Mbps.

11.4.2.2 Estimate Cell Capacity

As described in Chapter 7, cell capacity can be estimated based on simulations or analytical formulas. User data rate, traffic characteristics (variability, burstiness), quality of service requirements (delay, BER/FER), and outage probability are the main factors that determine the supported capacity. The higher the bit rate, the less users can be supported. The smaller the interference outage probability (i.e., the better the network quality), the smaller the provided capacity. It can be easily understood that provision for better quality radio connections requires more radio resources, and thus fewer users can be supported within a fixed amount of spectrum.

A possible impact of sectorization or adaptive beamforming, also called spatial division multiple access (SDMA), on spectrum efficiency needs to be estimated. It depends on specific radio environment and antenna equipment. For example, in microcells the signals tend to propagate along the street corridors no matter to what direction they were originally transmitted. Thus, the maximum sectorization in the downlink gain is less than in a more open space environment.

A CDMA radio network cannot be operated at a so-called *pole capacity*. Pole capacity can be defined as the theoretical maximum capacity. The *load factor* defines how close to the maximum capacity the network can operate. When determining a suitable load factor, a number of factors need to be taken into account. These include traffic characteristics, radio resource management algorithms, and even the planning capabilities of the network operator. Loads over 75% have been found to cause instability in the system [4].

Example 1: Assume 64-Kbps data service. Based on the network simulations performed in Chapter 7, we conclude that with full load the spectrum efficiency is 100 Kbps/MHz/cell (i.e., 500 Kbps/cell given 5-MHz bandwidth). Assuming 50% load, the actual spectrum efficiency is 250 Kbps/cell (i.e., slightly under 4 Erlang).

Example 2: Assume 8-Kbps speech users. Based on the network simulations performed in Chapter 7, we get spectrum efficiency of 108 kbps/MHz/cell (i.e.; 33.75 Erlang with 5-MHz bandwidth and 75% load). Assuming each user generates 30 mErlang traffic, one cell can support 1125 users.

11.4.2.3 Estimate Maximum Cell Coverage

The link budget is calculated according to principles presented in Chapter 7. In addition to the basic assumptions such as data rate and E_b/N_0 performance, the equipment-specific factors such as cable losses, antenna gain, and receiver noise figure need be taken into account.

Soft handover gain has a large impact on link budget. The soft handover gain is the gain brought by handoff at the boundary between two or more cells, where there is equal average loss to each of the cells. The soft handover gain can be obtained by first calculating the log-normal fade margin (i.e., the margin required to provide the specified coverage probability at the border of a single isolated cell, and then the

corresponding margin required at the boundary between two or more cells). The soft handover gain is given by the difference (in decibels) between the two different margins. The soft handover gain depends on the shadowing correlation and coverage probability. The larger the coverage probability requirement (i.e., smaller outage probability), the larger the required margin. In addition to the gain brought by cell selection, soft handover brings the macro diversity gain through increased diversity (see Section 7.2.5). The actual gain depends on the radio environment, and number of RAKE fingers.

Since each radio environment has its own characteristics, for more detailed coverage prediction some correction factors for the pathloss models presented in Chapter 4 are required. Field measurements can also be used.

For the uplink, the impact of load factor η in the link budget, the interference margin I_m (in decibels), can be determined from

$$I_m = 10\log\left(\frac{1}{1-\eta}\right) \tag{11.4}$$

and is illustrated in Figure 11.5. As can be seen, the interference margin increases, and thus, the range would decrease with increasing load factor. As was discussed in Chapter 7, multiuser detection can reduce the impact of load factor on the cell range.

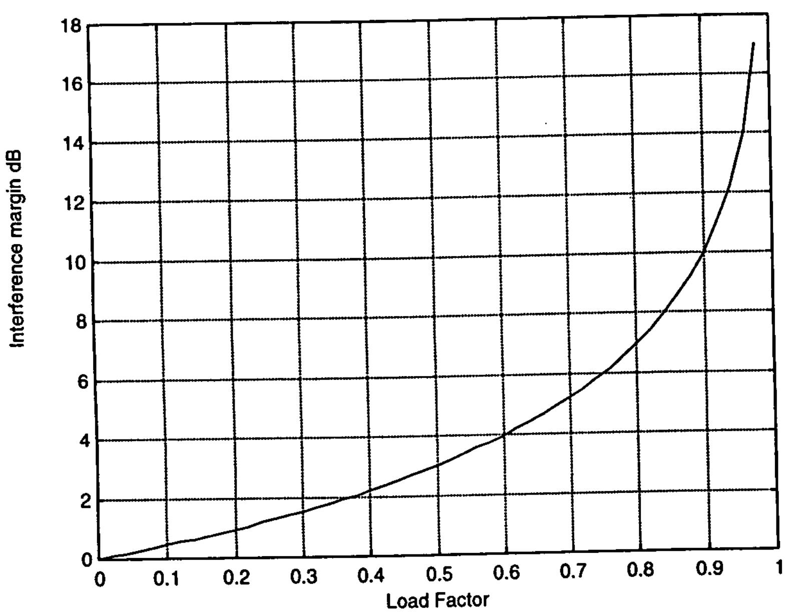


Figure 11.5 Interference margin as a function of load factor.

Asymmetric traffic has to be considered in the link budget calculations. CDMA can trade the uplink system load for coverage. This is useful since usually the mobile station transmission power limits the maximum cell range. If there is three times more downlink than uplink traffic, the uplink may only be loaded to 25%, when the downlink is loaded to 75%. Consequently, the uplink coverage will improve by 4.8 dB compared to the downlink. However, the downlink coverage can be compensated for by increasing the transmission power of the base station.

In Table 11.1, an example of link budget calculation is given. See Chapter 7 for a detailed explanation how to calculate the link budget. Note that in this calculation the carrier frequency and antenna height have been taken into account. We see that the uplink and downlink have different link budgets and thus different ranges. In the last column, we have different bit rates (144 and 28.8 Kbps) and unequal loading in the uplink and downlink. In practice, the downlink link budget is matched to the uplink by adjusting the base station transmission power. This has been done in the last case by increasing the downlink transmission power.

11.4.2.4 Estimate Cell Count

The number of cells required to cover a given area can be calculated based on the capacity and link budget. A network can either be *coverage* or *capacity limited.* Capacity limited means that the maximum cell radius cannot support the total offered traffic. In this case, the cell count is obtained as follows. By dividing the number of users, one cell can support by the number of user/km^2, we obtain the cell area in square kilometer. The total number of cells is obtained by dividing the total area by the cell area. Coverage limited means that there is enough capacity in a cell to support all traffic (i.e., the cell size could be larger from a capacity point of view, but the maximum cell range limits it). In this case, the maximum cell area is used to find out the required number of base stations.

Example: The area to be covered is 100 km^2 and the number of speech users is 5000/km^2. The cell area 23.32 km^2 is obtained from Table 11.1. One cell can support 1041 users with a 50% load, corresponding to 3-dB interference margin. Thus, the cell area is 1041/5000 = 0.2082 km^2 and altogether 100/0.2082 ≈ 481 cells are required to cover the area. If the user density was 30 user/km^2, then the network would be coverage limited since one cell could cover 37.4 km^2, from the capacity point of view. However, this exceeds the maximum cell area of 23.32 km^2. The required number of cell sites would in this case be 100/23.32 ≈ 5.

11.4.3 Detailed Network Planning

After the cell count has been obtained, detailed radio network planning, taking in account the exact radio environment where each cell is deployed, can start. Due to cost reasons, zoning laws, building restrictions, or some other impediment, it is not possible to obtain optimum cell sites in a real network. This will impact the initial coverage plan. For the detailed network planning a network planning tool is used. The previously

described process of obtaining rough cell count could also be included in a network planning tool. A network planning tool has a digital map of the area to be planned. Building heights and antenna patterns are also modeled. Since each vendor has its own tools with slightly different capabilities, we try to reflect here the general principles and not the detailed procedures used in real assessment. The optimization process of the radio network coverage includes:

- Detailed characterization of the radio environment;
- Control channel power planning;
- Soft handover parameter planning;
- Interfrequency handover planning;
- Iterative network coverage analysis;
- Radio network testing.

Table 11.1
Link Budget Calculation

Service	Speech		Data		Data		
	Downlink	Uplink	Downlink	Uplink	Downlink	Uplink	Unit
Average TX power/TCH	30	24	30	24	30	24	dBm
Number of code channels	1	1	1	1	1	1	
TX power/code channel	30.00	24.00	30.00	24.00	37.14	24.00	dBm
Max TX power/TCH	30.00	24.00	30.00	24.00	37.14	24.00	dBm
Max total TX power	30.00	24.00	30.00	24.00	37.14	24.00	dBm
Cable etc losses	2	0	2	0	2	0	dB
Antenna gain	13	0	13	0	13	0	dBi
TX EIRP/TCH	41.00	24.00	41.00	24.00	48.14	24.00	dBm
Total TX EIRP	41.00	24.00	41.00	24.00	48.14	24.00	dBm
RX antenna gain	0	13	0	13	0	13	dB
Cable and connector losses	0	2	0	2	0	2	dB
Receiver noise figure	5	5	5	5	5	5	dB
Thermal noise density	-174	-174	-174	-174	-174	-174	dBm/Hz
Interference margin	3	3	3	3	6	1.25	dB
Tot. noise + interf. density, No+Io	-166	-166	-166	-166	-163	-167.75	dBm/Hz
Information rate	8	8	144	144	144	28.8	kHz
10log(Information rate Rb)	39.03	39.03	51.58	51.58	51.58	44.59	dBHz
Eb/No (incl. macro div. gain)	8	6.6	8	6.6	8	6.6	dB
Receiver sensitivity	-119.0	-120.4	-106.4	-107.8	-103.4	-116.6	dBm
Handoff gain (selection gain)	5	5	5	5	5	5	dB
Expl. diversity gain	0	0	0	0	0	0	dB
Other gains / losses	0	0	0	0	0	0	dB
Body loss	3	3	3	3	3	3	dB
Log normal fade margin	11.3	11.3	11.3	11.3	11.3	11.3	dB
Maximum path loss	150.67	146.07	138.12	133.52	142.26	142.26	dB
Carrier frequency f	2000.00	2000.00	2000.00	2000.00	2000.00	2000.00	MHz
Base station antenna height, hB	15.0	15.0	15.0	15.0	15.0	15.0	m
Range	**3.97**	**3.00**	**1.84**	**1.39**	**2.37**	**2.37**	**km**
Cell area	**40.96**	**23.32**	**8.80**	**5.01**	**14.62**	**14.62**	**km^2**

11.4.3.1 Detailed Characterization of the Radio Environment

For optimization purposes, the radio environment has to be characterized in detail by actual field measurements or other techniques such as ray tracing [2]. Characteristics such as pathloss, shadowing, and delay spread are studied. Furthermore, a RF survey can be performed using spectrum analysis. RF measurements reveal competing cellular operators and background interference and possible intermodulation products.

11.4.3.2 Control Channel Power Planning

The transmission of control channels reduces the overall capacity of the network. The total transmission power depends on how the power is divided between pilot, synchronization, paging, broadcast, and traffic channels. Power allocation principles are related to coverage.

In FMA2 and wideband IS-95, a common pilot channel is used for handover measurements and coherent detection (see Chapter 6). In Core-A, a BCCH channel is used for handover measurements only (see Chapter 6). The coverage of these channels must be larger, when compared to traffic channels, in order for the mobile station to be able to decode other base stations before entering the soft/softer handover zone. Roughly 5% of total base station power will be allocated to the synchronization channel. Since the broadcast channel including the cell information has to be decoded before the mobile station enters the coverage area of a cell, for example 3 dB more power than for a 8-Kbps speech traffic channel should be allocated. The paging channel(s) can manage with about the same power as a 8-Kbps speech traffic channel.

11.4.3.3 Pilot Pollution

By pilot pollution it is meant that there are a number of pilot signals but none of them are dominant enough to allow the mobile station to start a call. The symptoms of such a situation are typically good mobile received power, poor E_c/I_0, and poor forward BER. To prevent pilot pollution, network planning needs to create a cell plan where a dominant pilot exists. This can be done by scaling pilot powers, down-tilting antennas, or increasing coverage of certain sectors or cells.

11.4.3.4 Soft Handover Parameter Planning

Soft handover performance impacts the required fade margin against shadowing and the number of users in soft handover. As was shown earlier, the fade margin is one factor in the link budget and thus will influence the number of base stations that have to be deployed. Users in soft handover require additional channel elements and backhaul connections (see Section 11.8). The network operator's goal is to minimize the fade margin and number of users in soft handover, while maintaining a satisfactory quality of service.

The number of users depends on handover thresholds and service type. It is possible that packet services do not use soft handover at all. Also, antenna tilt and orientation impact the percentage of cell coverage in soft handoff and should be considered in the handover parameter and region planning. The optimum active set size depends on the handover thresholds, number of available RAKE fingers, and radio environment. Typically, more than three base stations in the active set do not give significant soft handover gain increase.

In the following we review the soft handover planning based on the IS-95 soft handover algorithm. The main parameters for soft handover are handover thresholds, handover timers, and pilot search window at the mobile station. The parameters used in the active set updating are the following:

- T_{ADD}: this threshold indicates the point when a pilot should be added to the candidate set. The mobile station is measuring the pilot E_c/I_0.
- T_{DROP}: this threshold indicates when a pilot should be dropped from the active set if pilot strength has dropped below T_{DROP} for T_{TDROP} seconds.
- T_{COMP}: the mobile station reports that a candidate set pilot is stronger than an active set pilot only if the difference between their respective strengths is at least $T_{COMP} \times 0.5$ dB.

A typical value of T_{ADD} is between −12 and −16 dB [5]. Increasing the T_{ADD} threshold reduces the number of users in the soft handover and thus reduces the required equipment overhead. However, reducing it too much increases interference and thus reduces performance [5]. Since the time required to detect a new pilot should be minimized, filtering in the searcher for pilot in neighboring sets should be minimized. The T_{ADD} threshold should be high enough to prevent false alarms due to large noise in pilot measurements.

T_{DROP} should be low enough to prevent the loss of a good pilot. Typical values for T_{DROP} are between −17 and −20 dB. If the T_{TDROP} timer is too short, unnecessary handovers might occur, so T_{TDROP} should be in the order of the time required to establish handover. If T_{ADD} and T_{DROP} are too close to each other and the T_{TDROP} timer is very short, a handover ping-pong effect (i.e., adding and dropping the same pilot consecutively) might occur.

T_{COMP} should be set to a value that would prevent the mobile station from continuously sending active set update messages as a consequence of small changes in the strengths of pilots in the active set and the candidate set. Too large values, however, would introduce too long a delay before a pilot strength measurement message is issued, delaying the handover procedure [6].

The impact of T_{COMP}, pathloss slope and standard deviation of shadowing (log-normal fading), and different radio environments have been studied in [7]. The following results were reported: the higher the log-normal standard deviation, the higher the shadow margin required to combat fading. Higher pathloss slope causes better isolation between adjacent cells and hence a lower margin is required. T_{COMP} did not impact performance very much in [7]. The required fade margin increases with delay in establishing handover. A typical handover delay in today's systems is 2 to 5 seconds [8].

In the future, other criteria such as uplink interference might be required in the handover process. Furthermore, different services with different quality of service requirements might require different criteria for handover and even different active set sizes. This increases the network planning complexity.

Since soft handover gain depends on the correlation between the signal from different cell sites and on the correlation length of the shadow fading from the same cell site, field measurements are required to quantify the exact gain and thus the network capacity. Lower correlation between sites leads to higher call reliability. Furthermore, a large correlation length of the log-normal fading increases the fading margin [8]. The impact of correlation length depends on the mobile speed: the faster the mobile moves, the smaller the impact, since the correlation coefficient for the log-normal fading is smaller.

The searcher window determines the number of chip delays for which a mobile is instructed to look for pilot. A large window helps to collect all the multipath energy but leads to a slow measurement process, which might be detrimental for high-speed mobiles. A too narrow window leaves some of the significant multipaths unnoticed and thus degrades performance. Obviously, the window size is a function of delay spread and should be optimized for each cell, taking the expected mobile speed distribution into account.

11.4.3.5 *Interfrequency Handover*

The compressed mode is one alternative for seamless interfrequency handover (see Chapter 5). The main parameter of interfrequency handover is the frequency at which the base station commands the mobile station to measure neighboring carriers. The stronger the serving cell, the more seldom the mobile would check the neighboring channels. At cell edge, the mobile should measure pilot channels more often in order to be prepared for interfrequency handover.

The use of slotted frames leads to performance degradation. This degradation depends on the number of slotted frames. Assuming that the performance of a slotted frame is 2 dB worse compared to continuous frame, the overall degradation depends on the number of slotted frames. For example, if every 10th frame is slotted, the degradation would be 0.2 dB.

In case there are both micro- and macrocell coverage, then the criteria for which each layer is selected has to be decided. In general, if a fast moving mobile is detected in microcells it should be handed over to macrocells. Therefore, the speed of mobile stations needs to estimated, for example, from the fading characteristics. On the other hand, a slow moving mobile should be handed over to microcells. Wrong decisions should be avoided; for example, a mobile station in a car that has stopped at traffic lights should be kept in macrocells since it will speed up in the next moment.

11.4.3.6 *PN-Offset Planning*

In IS-95 and cdma2000, pilot signals belonging to different base stations and sectors are distinguished by the use of different phase offsets of the same pilot pseudo-noise (PN) sequences. If a propagation delay between two base stations exceeds the pilot signal offset, the so-called PN confusion occurs. PN confusion might cause dropped calls due to high interference from another base station or a handover to a wrong target cell. Therefore, PN-offsets need to be planned [9,10].

The distance d a pilot signal travels can be converted to a distance in chips using the following formula [10]:

$$T_d\,(\text{chips}) = \frac{B}{c} \times d \tag{11.5}$$

where B is the chip rate of the system and c is the speed of light (3×10^8 km/s). If the chip rate is 1.2288 Mcps as in IS-95 and the PN-offset index is 64 chips, one ON-offset index corresponds to a 15.6-km propagation distance. For cdma2000, one PN-offset index would correspond to 5.6 km. Figure 11.6 presents a scenario where the propagation distance difference of cell I and cell J is equal to PN-offset, therefore causing PN-confusion.

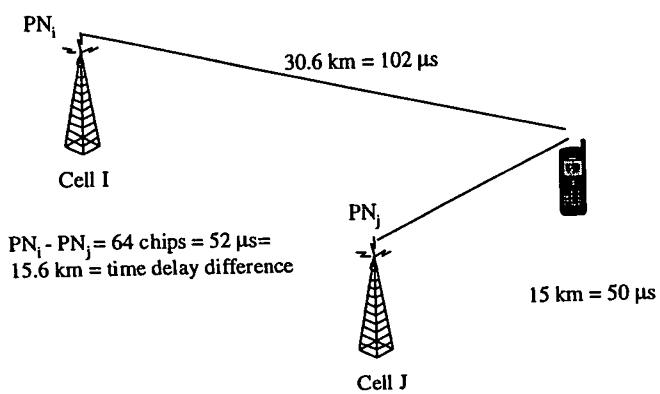


Figure 11.6 PN-confusion (*After:* [10]).

11.4.3.7 *Iterative Network Coverage Analysis*

Once the radio network environment has been characterized, control channel power allocation decided, and handover parameters planned, a detailed coverage analysis can be carried out. The ratio of in-cell to total interference is also referred to as the *frequency reuse coefficient* (see Chapter 7). The frequency reuse coefficient is unique for each cell. Starting from the initial configuration based on the preceding steps in the planning process (cell count, detailed characterization of the radio environment, control channel power planning, soft handover parameter planning, interfrequency handover planning), network quality is analyzed and new frequency reuse coefficients evaluated [5]. These are then used again to predict the coverage within different cells. This iterative process is repeated until convergence is achieved. A network planning tool can be used to automate the process. With this type of tool, quality plots can be created and gaps or holes in the coverage can be detected.

11.4.3.8 *Nonuniform Traffic*

In general, nonuniform traffic degrades the overall performance of the system. On one hand, the quality becomes poor due to increased interference in dense traffic zones; on the other hand, the quality becomes excessive. This dispersion of the communication quality restricts the system capacity [11]. The system efficiency can be improved by adaptive control of cell radius, antenna directivity, and uplink received power threshold. The cell radius is controlled by adapting the pilot transmission power. If the observed SIR is higher than required, the cell radius can be extended; and if it is lower, the cell radius is reduced. The desired received power threshold in the uplink is increased and reduced, respectively, to balance the cell radii of uplink and downlink. In a sector cell configuration, the central angle of each sector (i.e., the antenna directivity) is changed to equalize the communication quality in every sector belonging to the BS [11].

Of course, dynamic control of cell radius and antenna directivity requires careful planning and stable control system to prevent undesired effects. Also, interaction with handovers should be considered.

11.4.3.9 *Radio Network Testing*

After the network has been deployed, it is tested by drive tests to find out the network quality in practice. Furthermore, drive tests are used to collect data for optimization of the network performance. Before measuring, the optimization scope is determined so that correct measurements can be performed. Typical topics for measurements are coverage limitations, drop-call rate, quality, and pilot sets for soft handover.

Once the network has been set up, the work does not end. Testing continues as an essential part of the network operations. Therefore, proper operation and maintenance facilities are required to perform measurements and to collect data about network behavior. The measurement data are used to monitor the network quality and to locate congestion and quality gaps.

11.5 MICROCELL NETWORK PLANNING IN CDMA

Due to propagation environment and network topology, microcell deployment differs significantly from the macrocell deployment (see Chapter 4). Microcell base stations are installed at lamp post level and therefore signals tend to propagate along street canyons. Two critical aspects for microcell network planning are corner effect and soft handover design.

11.5.1 Corner Effect

Figure 11.7 illustrates the so-called *corner effect.* When a mobile moves around a corner, changes in received signal level at the mobile station happen very rapidly. In case there is a new base station behind the corner, the signal strength received by the mobile station rises very fast. If the mobile station could not acquire the new base station fast enough, the increased interference leads to a dropped call. On the other hand, since the new BTS cannot regulate the power of the mobile station, the high transmission power of the mobile station can block all users in the new cell. To reduce the impact of corner effect, a fast forward handover can be performed if the old base station is dropped. Corner effect can also be avoided by proper planning of cell locations and handover thresholds. Handover thresholds can be specified in such a way that the mobile station is in soft handover before and after a street corner (i.e., overlapping cells).

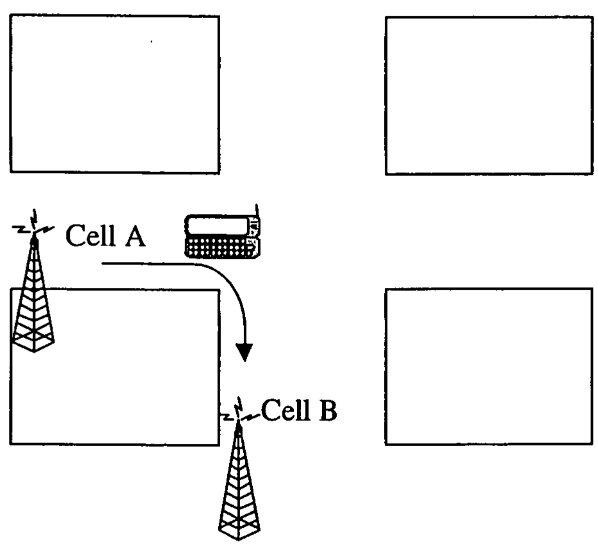


Figure 11.7 Corner effect.

11.5.1.1 Soft Handover Design in Microcell Environment

Since microcells are small, frequent handovers occur as the users move along the streets. The danger is that if traditional cell planning is used, then fast moving mobiles cause so much interference due to too slow handovers that capacity in microcells will drastically decrease. Furthermore, the signaling load increases considerably due to a large number of base stations. Thus, it is advantageous to deploy intelligence to the base stations.

The need for soft handovers in microcellular CDMA can be reduced by using distributed antenna and sectored cells. With distributed antennas, as shown in Figure 11.8, large coverage areas are achieved without use of many small micro- or picocells. The forward link capacity is not good, since the signal is transmitted to the whole coverage area even though the mobile is only in one location.

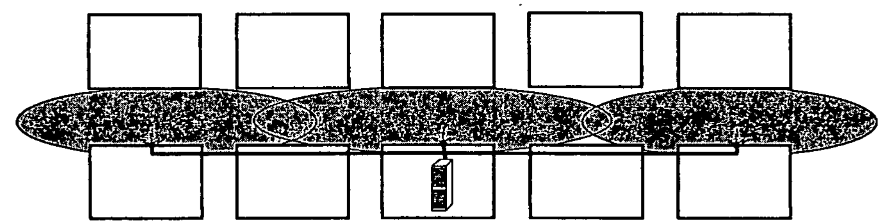


Figure 11.8 Street microcell with distributed antennas.

11.5.2 Micro/Macrocells in the Same Frequency

Placing micro- and macrocells in the same frequency is possible with low antenna installation and antenna tilting [12]. The danger is that the whole cell is in soft handover and no extra capacity is gained. In this solution, a fast soft handover from macrocell to microcell is required.

11.6 INDOOR PLANNING

The building types, sizes, construction, and materials exhibit a very large variation. Thus, no general propagation model that would be valid everywhere exists. Attenuation due to walls and floors depends on the construction of the building. Furthermore, propagation loss between floors depends also on structures like stairs and elevators. Therefore, field measurements or prediction of propagation by Ray-tracing are required to characterize individual buildings.

Due to short propagation distances, the values of indoor delay spreads are very small. Thus, the coherence bandwidth is large and consequently diversity order is low. Therefore, extra diversity may be added with delayed transmission or reception. One possibility is to use distributed antennas. By introducing time delay elements between the antennas in the distributed antenna system, deliberate multipaths are created that can

be processed by the CDMA RAKE receiver [13]. However, creating too many multipaths actually degrades the performance due to uncaptured energy.

The deployment of indoor cells into the same frequency as outdoor cells presents some challenges for the system operation. Since mobiles have to be connected to a cell where they use minimum transmission power, outdoor users can occasionally get the lowest power with picocells, for example, through windows. This situation most likely will not last very long. How is the handover working in this situation? One way to avoid this kind of situation is to deploy indoor cells to a different frequency band. Another possibility is careful network planning to avoid this. This means low transmission powers indoors and antennas with their backs towards windows.

11.7 SECTORIZATION AND SMART ANTENNAS

Sectorization increases the capacity of a CDMA system. In an ideal case, perfect isolation between sectors is achieved and the capacity will increase in direct proportion to the number of sectors. Of course, in practice, isolation is not perfect and the capacity increase is smaller. Furthermore, the number of softer handovers increases with the number of sectors. Several factors impact the sectorization gain. These include interference distribution and antenna location. Obstacles and structures near the antenna tend to change the antennas impedance and radiation pattern (see also Section 7.3.2.8). This causes the coverage to deviate from the planned coverage. Another impact is that coupling between diversity antennas might increase, and consequently, correlation increases, degrading performance. Usually it is assumed that 3 sectors result to a capacity increase of 2.5.

Smart antennas (also called SDMA) are defined as multibeam or adaptive array antennas without handover between beams [14]. Multibeam antenna uses multiple fixed beams in a sector, while in an adaptive array the received signals by the multiple antennas are weighted and combined to maximize the SNR. The advantage of antenna arrays compared to fixed beam antennas that in addition to the M-fold antenna gain they provide M-fold diversity gain. However, they require a receiver for each antenna, and tracking the antenna weights at the rate of the fading.

The M-fold antenna gain will increase the range by a factor of $M^{1/\gamma}$ where γ is the pathloss exponent, and reduces the number of base stations to cover a given area by $M^{2/\gamma}$ [14]. A multibeam antenna with M beams can increase the capacity by a factor of M by reducing the number of interferers. Adaptive arrays can provide some additional gain by suppressing interferers further. However, since there are so many interferers the additional gain might not be worth the complexity.

11.8 NETWORK DIMENSIONING

11.8.1 BTS Channel Element Planning

The number of channel elements can be calculated when the offered traffic in Erlang, control channel requirements, and number of users in soft handover are known. Calculations are based on Erlang tables that give the number of required channels for a given blocking probability and offered traffic.

In addition to the traffic channels that can be determined from capacity calculations, common control channels in the downlink and access channels in the uplink need channel elements. Each wideband CDMA carrier requires one pilot and at least one common control channel, which consists of the paging channel, synchronization channel, and broadcast channel. In the uplink, the number of access channels depends on the expected access load and synchronization time requirements. In case an umbrella cell is used to overlay microcells, the microcells do not necessarily need that many common control channels and access channels if the initial access traffic is handled mainly through macrocell. However, for packet access, common control channels are required also in microcells.

The additional number of channel elements due to soft handover depends on the soft handover type. Softer handover does not need an additional channel element since the same channel element can handle signals from different sectors within the base station. Two-way soft handover needs two channel elements, and three-way soft handover needs three channel elements. As was discussed previously, the number of users in soft handover depends on the radio environment, handover thresholds, and antenna configuration and it can be found out from simulations or field measurements.

Due to the trunking gain effect, it is beneficial to pool channel elements between sectors and different carriers in the same cell site. This also applies to the channel elements due to soft handover. Therefore, the actual number of additional channel elements is less than the percentage of users in soft handover due to improved Erlang efficiency as will be illustrated by the following example.

Example: Table 11.2 shows a calculation for the number of required BTS channel elements. We consider a reference BTS implementation, which has one CDMA frequency and three sectors and is realized as one cluster with a common pool of "floating" channel elements. The total carried traffic per sector is 39 Erlang including soft handover overhead traffic of 9 Erlang. Thus, the total traffic for the cell site is 3×39 = 117 Erlang. For a blocking rate of 2%, the total amount of needed channel elements for the cluster is 130, or 43.3 per sector (using the Erlang-B formula applied for 117 Erlang). In addition, for the downlink one channel element for pilot channel and for common control channel is required. In the uplink, one channel element for access channel is assumed. Thus, the overall number of channel elements per sector is 46.

Without additional soft handover traffic, the required number of channel elements for traffic channels would be 34.3. The additional number of channel elements

due to soft handover is 26.2%, which is less than the 30% overhead traffic and is due to the Erlang effect.

If there was no pooling between sectors, the required number of channel elements for traffic channels would be 49 (Erlang-B formula applied for 39 Erlang). Thus, the pooling reduces the required number of channel elements by 12%.

For high bit rate services, we have 8 Erlang traffic. Assuming 30% overhead traffic due to soft handover, we have 10.4 Erlang per sector and 31.2 Erlang per site, 41 channel elements = 41/3 = 13.7. The additional soft handover traffic 33/3=11 is 8 channel elements (i.e., 24.2%).

Table 11.2
Channel Element Calculation

	Speech 8 Kbps	Data 144 Kbps
Erlang/sector	30 Erl	8 Erl
Soft handover overhead	30%	30%
Total traffic (3 sectors)	117 Erl	31.2 Erl
Number of channel elements/sector from Erlang B table, pooling taken into account	44	14
Channel elements for pilot and common control channels	2	2
Total number of channel elements	46	16

11.8.2 Number of BSCs and Switches, HLR and VLR Signaling Traffic

The number of BSCs is determined based on the maximum configuration. Different vendors support different traffic loads through the BSC. Similar to BSC, switches from different vendors support different traffic loads, which also determines the number of required switches.

In order to determine signaling related to mobility management (i.e, between HLR and VLR), mobility models are required. The mobility models describe the mobility of wireless users (see Section 4.7).

11.8.3 Transmission Capacity

The required transmission capacity depends on the user and signaling data that needs to be transmitted between network elements. Connections include at least BTS-BSC, BSC-switch, and possibly BSC-BSC connections for handover purposes. Transmission capacity depends on the transmission technology. ATM, for example, provides good multiplexing gain. Soft handover traffic increases the required transmission capacity up to the network element, most likely to the BSC, where soft handover is terminated. In addition, soft handover between cells belonging to different BSCs over switches might

require special arrangements. It also has to be considered whether additional blocking is allowed for transmission.

11.8.4 Transmission Network Optimization

When the rough network dimensioning has been obtained, an optimization phase is carried out. Depending on the costs of transmission, different network configurations can produce optimum overall cost. Sometimes it is beneficial to have more BSCs than required to tailor the transmission network for the most optimum configuration. If a BTS is situated in the middle of two BSCs, then routing of that BTS needs to be determined based on transmission costs that might be different for the two cells.

11.9 CO-EXISTENCE

When a wideband CDMA network is deployed over a given area, other networks will nevertheless continue to operate. Most likely the W-CDMA network will co-exist with GSM, IS-95, and PDC systems. Two main scenarios for the deployment of wideband CDMA are then predictable:

- Wideband CDMA introduced in third generation frequency band;
- Wideband CDMA introduced into the second generation frequency bands.

An existing cellular operator will most likely start UMTS with local coverage, relying on the second generation system for low bit rate wide area coverage. Therefore, it is wise to ease the dual-mode terminal design by properly selected air interface parameters. In the longer term, the replacement of a second generation system by wideband CDMA depends on the gained advantages, such as higher bit rates or increased spectrum efficiency. Furthermore, licensing and availability of spectrum also impact the decision. If the available bands of an operator are congested, the higher spectrum efficiency of wideband CDMA might allow the operator to pack existing users into less spectrum and to create room for higher rate services. However, the introduction of a wideband carrier into congested frequency bands is fairly difficult. First, the operator needs to preload the network with dual-mode terminals with new wideband capabilities. Next, he needs to release the spectrum from old carriers and switch the new carrier on almost simultaneously to avoid interruption in service.

A Greenfield third generation operator in a country with an existing cellular network also needs dual-mode terminals. Only if the data market emerges very fast and justifies the investment in a new nationwide network can an operator rely on single-mode terminals.

The transition to third generation systems depends on technical possibilities such as increased spectrum efficiency and data rates but most of all on market needs and regulatory conditions. In this section, we highlight some of technical aspects related to network planning facilitating smooth transition and co-existence. The transition from the first generation technology AMPS to the second generation digital technologies in

the United States gives us some background for the transition to third generation as well. To avoid problems with co-existence of different air interfaces, the following aspects need to be considered from the network planning perspective:

- Intermodulation;
- Deployment scenarios;
- Guard zones and bands;
- Transition aspects;
- Handover between systems.

11.9.1 Intermodulation (IM)

A power amplifier is nonlinear near the saturation point. In this nonlinear area, the signal is distorted and produces harmonics. If two or more different frequencies are present at the input of an amplifier, intermodulation products are generated [15]. Figure 11.9 illustrates intermodulation power and its relation to the power of the original signal. As can be seen, the power of the third order intermodulation product increases three times faster than the power of the original signal. Equally, the power of the fifth order intermodulation product increases five times faster.

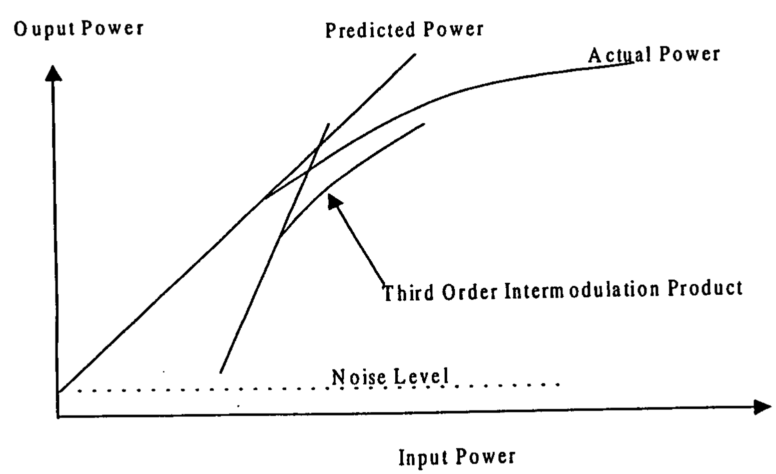


Figure 11.9 Intermodulation power (*After*: [15]).

In mobile radio systems, there are two types of intermodulation products: transmit IM and mobile generated IM [16]. The mobile station generated IM is more significant than transmit IM from the base station. Transmit IM is generated at the base

station transmitter. Several carriers are passed through the same power amplifier creating intermodulation products that might fall in the wideband CDMA band. Mobile generated IM is caused by the nonlinear effects in the active stages of the receiver front-end; these stages typically consist of a RF low noise amplifier (LNA), mixer and IF amplifier. As was discussed previously, the power of the third order intermodulation products, considered as the primary source for IM interference, is increasing very fast as the RF gain increases. Thus, when the mobile is close to the interfering base station, the third order products become severe in relation to the desired CDMA signal causing possibly a dropped call. One proposed solution to avoid this is to switch out the LNA to eliminate strong IM signals from occurring at the mobile station front-end prior to mixing [17]. However, if the desired signal is already weak, this might lead to a dropped call as well.

If wideband CDMA is deployed within the same frequency band as an existing network, there are two possibilities to map CDMA cells and existing cells. One alternative is *one-to-one mapping* of the cells of the wideband CDMA and the other system. Another alternative is that one wideband CDMA cell can cover more than one cell of the existing system (1-to-*N* mapping), since, in the beginning, wideband CDMA deployment load is still low and the network can be built to be coverage limited. Furthermore, the existing system has already reached quite small cell sizes due to large capacity demand. Later, when wideband CDMA capacity demand increases, it is likely that it will have either the same cell size as the other systems or an even smaller cell size.

From the deployment point of view, one-to-one mapping is better with respect to intermodulation product avoidance. There should be no intermodulation interference problems since wideband CDMA has similar power than the interfering system. Of course, this will lead to increased network infrastructure costs in the beginning. In 1-to-*N* mapping, the CDMA mobiles with small signal power might receive high intermodulation products. When a wideband CDMA mobile at the CDMA cell edge is at the sensitivity limit, then high power intermodulation products might block the mobile. This kind of deployment requires more planning.

11.9.2 Guard Bands and Zones

A second issue is frequency co-existence. Due to spectrum regrowth proper guard bands need to be allocated in order to not block the system by adjacent spectrum or to avoid that the other system blocks the wideband CDMA system.

A guard zone around the CDMA coverage area is needed to avoid co-channel interference. An example of an uncoordinated co-channel interference situation is when different technologies in the same frequency band are not separated well enough by a geographical barrier. Such a situation has been experienced in the United States between PCS market areas. The Code of Federal Regulations (CFR) Title 47 Part 24 addresses issues related to interference caused by the system. It specifies, for example, attenuation rules outside the operators' own spectrum (Section 24.238) and allowed field strength at the border of an operator's licensed service area (Section 24.236). The

National Spectrum Managers Association (NSMA) has prepared a document titled "Inter PCS Co-block Coordination Procedures" addressing the intersystem coordination rules for PCS systems. It specifies coordination distance, allowed degradation due to external interference, interference calculation procedures, and other coordination practices to prevent harmful interference from one system to another.

The guard zone depends on the interference that the system can tolerate from other systems and vice versa.

Example: For GSM the required carrier-to-interference ratio is

$$\frac{C_{\mathrm{GSM}}}{I} = 9\ \mathrm{dB} \tag{11.6}$$

We allow 1 dB degradation in C/I due to interference from CDMA:

$$\frac{C_{\mathrm{GSM}}}{I + I_{\mathrm{CDMA}}} = 8\ \mathrm{dB} \tag{11.7}$$

Thus

$$\frac{C_{\mathrm{GSM}}}{I_{\mathrm{CDMA}}} > 15\ \mathrm{dB} \tag{11.8}$$

From this we can now calculate the minimum separation between GSM and CDMA base stations using the same frequency. The ratio between cell size R and distance D can be calculated from [18]

$$q = \frac{D}{R} = \left(6\frac{C}{I}\right)^{1/\lambda} \tag{11.9}$$

Deployment within existing footprint reduces deployment efficiency. If coverage- limited deployment is desired, this might not be feasible since existing sites and their antennas are used. Furthermore, guard zones decrease the capacity of existing systems and depend on the deployment scenario.

Guard zones between IS-95 and AMPS have been studied in [19]. For a coverage limited system, they are larger in order to maintain the quality of the IS-95 network. Within a guard zone, the CDMA carrier plus guard bands cannot be used. So, even if IS-95 had larger capacity and coverage and it could be deployed using a lesser number of sites than AMPS, it proves to be that due to guard zone reasons, one has to convert enough sites to IS-95 to increase the total traffic in the network [19].

11.10 FREQUENCY SHARING

Frequency sharing means sharing the allocated frequency spectrum in the same geographical area between two or more operators. It is one potential technology to increase the overall capacity of a set of mobile radio networks in the same area. By sharing the frequencies, the trunking loss due to the subdivided total frequency band between operators is avoided. The frequency sharing is possible if the co-channel interference power of the interfering systems is less than the co-channel interference power from the own system. However, since CDMA systems are due to have a DCA scheme, frequency sharing might be possible only if the operators use the same cell sites.

REFERENCES

[1] Yang, S., *CDMA RF System Engineering*, Boston: Artech House, 1998.

[2] Cheung, J. C. S., M. A. Beach, and J. P. McGeehan, "Network Planning for Third Generation Mobile Radio Systems," *IEEE Communications Magazine*, Vol. 32, No. 11, November 1994, pp. 54–59.

[3] Jakes, W. C., *Microwave Mobile Communications*, New York: John Wiley & Sons, 1974.

[4] Stellakis, H., and A. Girdano, "CDMA Radio Planning and Network Simulation," *IEEE Int. Symbosium on Personal, Indoor and Mobile Radio Communications PIMRC'96*, Taiwan, 1996, pp. 1160–1162.

[5] Wallace, M., and R. Walton, "CDMA Radio Network Planning," *Proceedings of VTC'94*, Stockholm, Sweden, 1994, pp. 62–67.

[6] Qualcomm, *CDMA Network Engineering Handbook*, 1993.

[7] Chopra, M., K. Rohani, and J. D. Reed, "Analysis of CDMA Range Extension due to Soft Handoff," *Proceedings of VTC'95*, Chicago, IL, USA, May 1995, p. 917–921.

[8] Rege, K., S. Nanda, C. F. Weaver, and W.-C. Peng, "Analysis of Fade Margins for Soft and Hard Handoffs," *Proceedings of IEEE 6th Personal, Indoor and Mobile Radio Communications Conference (PIMRC'95)*, Toronto, Canada, 1995, pp. 829–834.

[9] Chang, C. R., J. Z. Van, and M. F. Yee, "PN Offset Planning Strategies for Non-Uniform CDMA Networks," *Proceedings of VTC'97*, Vol. 3, Phoenix, Arizona, USA, May 4-7, 1997 pp. 1543–1547.

[10] Yang, J., D. Bao, and M. Ali, "PN Offset Planning in IS-95 Based CDMA Systems," *Proceedings of VTC'97*, Vol.3, Phoenix, Arizona, USA, May 1997, pp. 1435–1439.

[11] Sato, S., and Y. Amezawa, "A Study on Dynamic Zone Control for CDMA Mobile Radio Communications," *Proceedings of ICC'97*, Vol. 2, Montreal, Quebec, Canada, June 8-12, 1997, pp. 984–988.

[12] Shapira, J., "Microcell Engineering in CDMA Cellular Networks," *IEEE Transactions on Vehicular Technology*, Vol. 43, No. 4, November 1994, pp. 817–825.

[13] Xia, H. H., A. B. Herrera, S. Kim, and F. S. Rico, "A CDMA-Distributed Antenna System for IN-building Personal Communications Services," *IEEE Journal on Selected Areas in Communications*, Vol. 14, No. 4, May 1996, pp. 644–650.

[14] Special Issue on Smart Antennas, *IEEE Personal Communications*, Vol. 5, No. 1, February 1998.

[15] Faruque, S., *Cellular Mobile Systems Engineering*, Boston: Artech House, 1996.

[16] Hamied, K., and G. Labedz, "AMPS Cell Transmitter Interference to CDMA Mobile Receiver," Proceedins of VTC'96, Vol. 3, Atlanta, Georgia, USA, April 1996, pp. 1467–1471.

[17] Gray, S. D., and T. Kenney, "A Technique for Detection AMPS Intermodulation Distortion in an IS-95 CDMA Mobile," *Proceedings of PIMRC'97*, Vol. 2, Helsinki, Finland, September 1-4, 1997, pp. 361–365.

[18] Lee, W. C. Y., *Mobile Cellular Telecommunications Analog and Digital Systems*, Second Edition, New York: McGraw-Hill, Inc., 1995.

[19] Ganesh, R., and V. O'Byrne, "Improving System Capacity of a Dual-Mode CDMA Network," *International Conference on Personal Wireless Communications*, Mumbai, India, December 1997, pp. 424–428.

Chapter 12

NETWORK ASPECTS

12.1 INTRODUCTION

In this chapter, we discuss the network aspects related to the development of third generation wideband CDMA systems. Since many of the network related aspects are independent of the radio interface, wideband CDMA itself is not discussed in depth in this section. The focus of this chapter is on the evolution from second generation networks towards third generation systems. It should be noted that most of the aspects described in this chapter are still under development in the standardization and thus are subject to change. This evolution proposal presents the author's view on the subject.

Section 12.2 describes a generic design methodology for mobile radio systems. This will guide the reader to understand the formal methodology used in the design of modern telecommunication systems. The different models used in the design methodology are defined. The open system interconnection (OSI) reference model and the Integrated Services Digital Network (ISDN) protocol reference model, as well as their application to the design of modern wireless systems, are introduced. The reference architecture defines the network elements and interfaces between them. The main network elements found in today's second generation systems are defined in the context of a network reference architecture. The already established functionality of second generation wireless systems is discussed with the help of functional planes. Finally, the preceding concepts are tied together by the stack and protocol architecture; the functions of different functional planes are mapped into the stack and protocol architecture, which is based on the OSI and ISDN protocol reference model principles.

The distinction of access and core network is one of the leading principles in third generation standardization work. Section 12.3 explains how this division impacts the network architecture of third generation systems. Furthermore, the Generic Radio Access Network (GRAN) and "family of systems" concepts defined in ETSI and ITU, respectively, are discussed.

In Section 12.4, several important technologies with respect to third generation

networks are discussed, namely intelligent networks (IN), asynchronous transfer mode (ATM), SS7 signaling system, and mobile application part (MAP) protocol.

Section 12.5 describes the new service capabilities required for third generation networks. Based on this, the new requirements that these new capabilities put on the current second generation networks are identified.

In Sections 12.5 and 12.6, the network architecture and protocols of GSM- and IS-41/IS-95-based systems are described. The impact of third generation requirements on these networks is identified and evolution possibilities are suggested.

While this chapter covers the evolution of the second generation mobile radio systems to third generation systems, the basic principles of evolution apply to any system. We try to highlight the procedure of evolution, which consists of the identification of new needs, in the assessment of current capabilities, and in the mapping of the evolution needs into interfaces and protocols of the existing systems and standards.

12.2 DESIGN METHODOLOGY

Modern communications systems are very complex collections of different functions, and mobile radio systems are no exception. In order to manage the design of such systems it is essential to use some structured modeling and design methodology. There exist several different system design methodologies, for example, the IN conceptual model and the CCITT 3-stage methodology.

The CCITT 3-stage methodology was originally developed for the design of ISDN-based networks [1]. Since wireless networks contain many new elements compared to ISDN-based fixed networks, the 3-stage methodology is usually applied in an ad hoc manner in the design of wireless telecommunication networks. A simplified view of the 3-stage methodology is illustrated in Figure 12.1. First, the new services planned to be offered by third generation networks need to be characterized. Based on this, new functional capabilities to support these services are identified. Finally, the functions are mapped into a reference architecture, which defines network elements and interfaces. For protocol design, the identified functions are mapped into a combined stack and protocol architecture based on OSI and ISDN reference models.

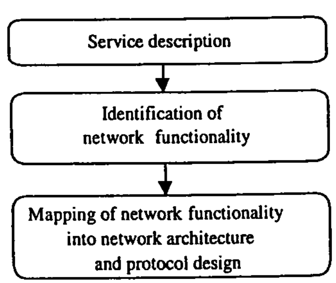


Figure 12.1 Design stages.

In the following five sections, the OSI and ISDN reference models, reference architecture, functional planes, and stack and protocol architectures are described.

12.2.1 OSI and ISDN Reference Model

The OSI model [2] shown in Figure 12.2 specifies seven different layers, each consisting of a set of functions that can be implemented independently from the functions in the other layers. The OSI model specifies communication principles between two network elements. The abstraction level of the model increases from bottom to top. This also means that in a time axis the span of events is shorter in lower layers and increases towards higher layers. In this way, communication between network elements can be defined in a modular manner. In addition to the peer entity, each layer communicates with the layers in the same network element immediately below and above it.

The physical layer of the OSI model is often referred as layer 1, link layer as layer 2, and network layer as layer 3. The rest of the layers are abstracted as higher layers. In the following, we use both naming conventions.

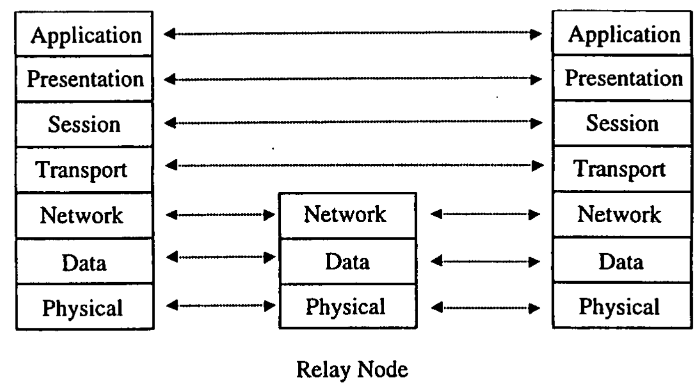


Figure 12.2 The OSI model.

The ISDN protocol reference model (PRM) is similar to the OSI reference model in the respect that it also organizes communication functions in layers and describes the relations of these layers with respect to each other [3]. In the modeling of wireless networks both models are used. The ISDN PRM introduces a division between the so-called *user* and *control planes*. The user plane consists of all functions being in charge of transferring user data, and the control plane consists of all functions being in charge of transferring information for the control of user plane data. Since usually at lower protocol layers a distinction of user and control data flows cannot be made, a third plane, *transport plane*, has been introduced. This is not explicitly described in the OSI reference model. It corresponds to the lower OSI layers (physical and data link and

possibly the network layer). Figure 12.3 shows the scope of user, control, and transport planes.

In the modeling of wireless networks, both OSI and ISDN PRM models are used. GSM is the first system that has followed these principles. The OSI model forms the basis for vertical protocol division, while the ISDN PRM model forms the basis for defining different protocols for user and control planes. Therefore, usually two different protocol stacks, the user and control plane, are defined. Sometimes the control plane is referred to as the *signaling plane*.

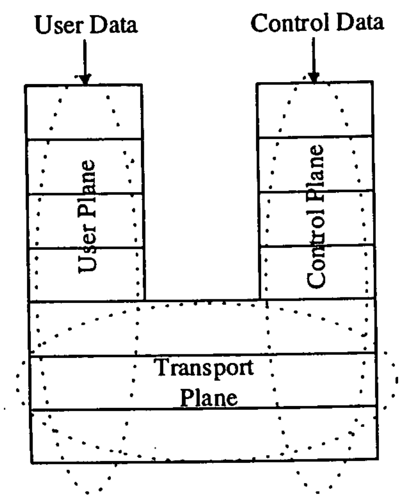


Figure 12.3 User, control, and transport planes in the OSI model.

12.2.2 Reference Architecture

The reference architecture specifies the network entities and the interfaces between these entities. Each network entity comprises a set of functions and is responsible for performing its allocated tasks. A sound architecture requires that related functions should be grouped in the same network entity. A typical mobile radio system contains at least the following network elements: mobile station, base station, and switch and data bases for roaming purposes such as home and visitor location registers (HLR and VLR). Figure 12.4 depicts an example of reference architecture. In the following, we depict the network entities found in digital mobile radio systems such as GSM, IS-95, IS-136, and PDC. These elements and interfaces (denoted by letters in Figure 12.4) will be then elaborated on in Sections 12.6.1 and 12.7.2 for GSM and IS-95/IS-41, respectively. The basis for the description is the GSM specification GSM 03.02 and the TR-45 reference model for IS-41.

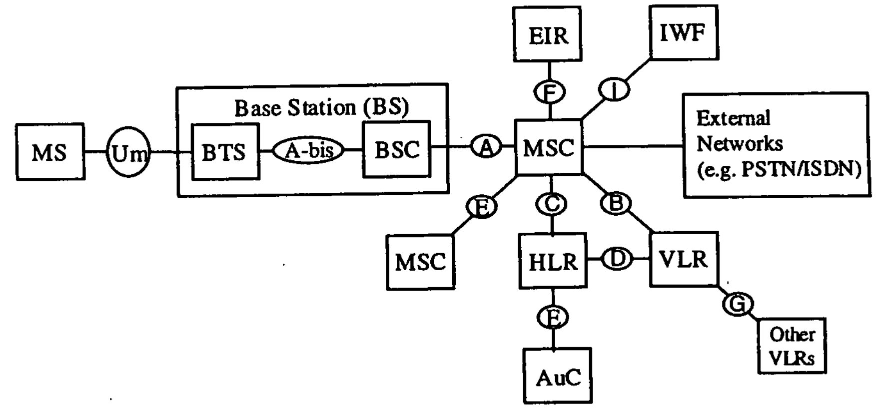


Figure 12.4 Example reference architecture.

12.2.3 Network Elements

In this section, different network elements in the reference model are described. The mobile cellular network itself is a public land mobile network (PLMN), which can connect either to external networks such as public switched telephone network (PSTN), ISDN, and public data network, or to another PLMN.

12.2.3.1 Mobile Station

The mobile station (MS) consists of the physical equipment used by a subscriber. The mobile station is usually divided into mobile termination (MT) and terminal equipment (TE) parts. MT terminates the radio path and contains functions such as modulation, error correction, and mobility management. TE is a terminal (e.g., data terminal) connected to a MT possibly via a terminal adaptation function.

12.2.3.2 Base Station System

The base station system (BSS) terminates the radio path in the network side. It communicates with the mobile station and connects to the mobile switching center through an interface termed the A-interface. Often, a so-called Abis-interface is implemented in such a way that the BSS consists of one base station controller (BSC) and one or more base transceiver stations (BTS). One BSC controls one or more BTSs. A BTS is a network component, which serves one cell.

12.2.3.3 *Mobile-services Switching Center*

The mobile-services switching center (MSC) is an exchange that performs all the switching and signaling functions for mobile stations located in a geographical area designated as the MSC area. The main difference between a MSC and an exchange in a fixed network is that the MSC has to take into account the mobile nature of the subscribers. Therefore, it has to perform procedures required for the location registration.

12.2.3.4 *Home Location Register*

This functional entity is a database in charge of the management of mobile subscribers. It maintains subscriber information (e.g., international mobile station identity, user profile). Furthermore, it contains some location information that enables the charging and routing of calls towards the MSC where the MS is located (e.g., the MS roaming number, the VLR address, the MSC address, the local MS identity). One home location register (HLR) can be associated with one or more MSCs.

12.2.3.5 *Visitor Location Register*

The visitor location register (VLR) is the location and management database for the mobile subscribers roaming in the area controlled by the associated MSC(s). Whenever the MSC needs data related to a given mobile station currently located in its area, it interrogates the VLR. When a roaming mobile station enters a new service area covered by an MSC, the MSC forwards a registration request (or a location update request) to the associated VLR, which informs/interrogates the HLR. Usually, one VLR is associated and co-located with one MSC.

12.2.3.6 *Authentication Center*

The authentication center (AC or AuC) is associated with an HLR, and it manages the authentication and encryption parameter for each individual subscriber.

12.2.3.7 *Equipment Identity Register*

The equipment identity register (EIR) is used for keeping track of mobile stations and their identities. Its purpose is to be able to record and prevent the use of possibly stolen or otherwise misused terminals.

12.2.3.8 *Interworking Function*

The interworking function (IWF) is a functional entity usually associated with the MSC. It allows interworking with external networks.

12.2.4 Functional Planes

Functional planes identify the functions performed in a cellular radio system to make the system work. The functions of each plane can then be mapped in the OSI model and the network architecture. This mapping serves as a basis for the detailed protocol and interface specification. In second generation systems, five generic functional planes are identified [4]:

- Communications management (CM);
- Mobility management (MM);
- Radio resource management (RRM);
- Transmission;
- Operations, administration, and maintenance (OAM).

The communication management layer includes functions related to setting up, maintaining, and releasing calls between users (i.e., call control (CC)). Furthermore, the CM layer has means to manage calls with so-called supplemental services. These include call forwarding, call waiting, and call hold. The third function of the CM layer is the management of short message services (SMS).

Mobility management covers the functions related to keep track of mobile subscribers within the home network and when they are roaming. Roaming is the use of services from other networks than the terminal's home network. Specific functions to manage these tasks are location update, paging, and cell and network selection. In order to obtain service from a cell, a user must register to the location area of that cell. This is called a location update. Before the network routes a call towards the subscriber, it pages the mobile station. The page is only sent to those cells belonging to the location area the mobile station currently belongs to. Obviously, the size of the location area is a trade-off between the signaling required for location updates and paging traffic. Roaming requires that there exist means to pass information between the home and visited networks. This is handled by the HLR and VLR, and signaling between them.

The task of radio resource management (RRM) layer is to maintain communication between terminals and base station. In order to do this, RRM performs the following tasks:

- Admission control;
- Channel assignment;
- Load control;
- Power control;
- Handover.

Admission control determines whether to accept the requested new communication, which can be based on several criteria. The channel assignment function assigns a physical channel for the connection. Load control manages the network load: it can reduce the traffic in the network, for example, by reducing the bit

rate for certain users. The radio resource management functions, including power control and handover, were discussed in Chapter 5.

The transmission layer is concerned with the transfer of signaling and user data. For the user data, the system needs to provide a consistent end-to-end transmission path, therefore, translation functions are required between different parts of the network. For example, currently the speech over the GSM air interface is transmitted with a 13-Kbps data stream. On the other hand, in the infrastructure side 64-Kbps PCM links are used. Thus, we need adaptation between these two rates.

The OAM plane contains functions related to monitoring and controlling of the system. These functions include billing and accounting, and collecting performance data.

12.2.5 Stack and Protocol Architecture

Now we have defined the concepts required for understanding the definition of a modern telecommunications network: reference architecture, OSI and ISDN reference model, and functional planes. To put these concepts together we need to map the functions identified in different functional planes into the OSI model layers and into the network elements. This is performed with the allocation of the identified functions[1] to a network/physical architecture (i.e., a set of network elements) described by *stack and protocol architecture* illustrated in Figure 12.5. As can be seen from Figure 12.5, several protocols can cross one interface, and some of the protocols exchange messages between nonadjacent network elements. An interface is a point of contact between two adjacent entities. In the stack architecture, the physical architecture is presented by means of protocol stacks. The functions are allocated to network elements and to protocol layers. It represents a "vertical" view of protocol layers. The protocol architecture focuses on certain protocol layers (i.e., it shows where a given protocol layer or sublayer is terminated). It represents a "horizontal" view with respect to protocols. In Figure 12.5, the different protocols are denoted with a letter P and two numbers indicating the layer number and the protocol in question.

The CM, MM, and RRM planes concern the control of the system. They are mapped into layer 3 of the OSI model. In order to perform their tasks, CM, MM, and RRM protocols need to exchange messages among peer entities located in different network entities and among other protocol entities. Thus, their specification consists of a definition of *messages* and the use of these messages.

The transmission functions within the mobile network, comprising of a MS, BS, MSC, HLR, VLR, and some additional network entities, can be mapped into three different OSI layers: physical layer (layer 1), link layer (layer 2), and network layer (layer 3). The end-to-end connection between two terminals also uses higher OSI layers; for example, for data connections the transmission control protocol (TCP) is used in the transport layer.

[1] A function can be described by a state machine or an algorithm. Network protocol designers consider a function as a state transition description (i.e., a state machine), while radio interface designers identify a function as an algorithm [5].

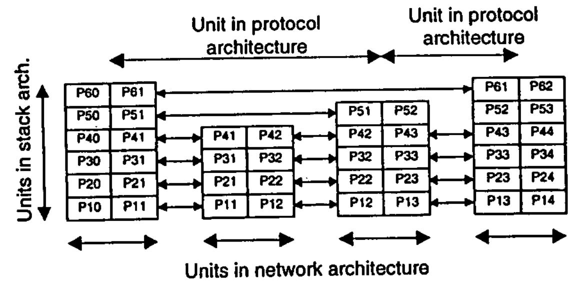


Figure 12.5 Stack and protocol architecture.

The link protocol structures the information to be transmitted in units bigger than a single bit. These units are called *frames*. When the maximum length of a message passed to the link layer exceeds the predefined maximum length, segmentation and reassembly have to be performed. The link layer also performs error detection and correction as well as quality monitoring and flow control.

While the link layer enables the exchange of information between directly interconnected network elements, the network layer provides additional transmission functions for information transmission with end-to-end connections through several network elements. The network layer covers aspects such as routing (i.e., how messages are passed from one point to another until they reach their final destination). Routing requires an addressing mechanism. For example, when a mobile station sends a message, it can go to the MSC, BSC, or HLR.

The above-described generic modeling based on application of the CCIT 3-stage model presents a simplified view of the actual system development. GSM is the first wireless system that has followed structured specification principles, and this is reflected in its sound architecture and protocols. However, even GSM has peculiarities where interfaces and protocols mix with each other and protocol boundaries are not clear. And sometimes, the conceptual distinction between protocols and interfaces is ill resolved in the actual GSM specifications [4].

The concepts of interface, protocol, and network elements will help us to understand how the second generation networks evolve towards the third generation. When new services such as multimedia calls are introduced, we need to first identify the new fuctionalities in the different planes required to support them. Then, these functions are mapped into protocols, interfaces, and network elements to see how they need to be developed. We might need a machine to fulfil a new function, for example the GPRS support node (GSN) in GSM (see Section 12.6.1), which was introduced in order to support packet switched services. On the other hand, the introduction of a new network element results in a new interface such as the Gb interface between BSS and GSN. An existing interface might also require to be changed to accommodate new services.

12.3 CORE AND ACCESS NETWORKS

When discussing third generation systems, the term core network and its evolution are frequently mentioned. A functional network architecture can be described with the concepts of access and core networks, as illustrated in Figure 12.6, in which the GSM network entities have been used as examples.

An access network comprises all functions that enable a user to access services. Furthermore, an access network can be used to hide all access-specific peculiarities from the core network. For example, in the case of a radio access network, all air interface–related functions should be kept within the access network part.

The core network comprises the switching network and service network. A switching network includes all the functions related to call and bearer control for fixed transmission. The MSC in GSM is an example of a switching network entity. A service network comprises all functionalities for the support services including location management.

Since most of the radio related parts are located in the access network, the core network should be little impacted by a new radio interface, and can evolve partly independently of the access network. Therefore, one of the drivers in third generation standardization, especially in Europe, has been a more clear separation between access and core network than that of the current systems. This is motivated by the isolation of radio-related functions from the switching functions. This separation has led to the concept of GRAN in ETSI, described in the next section.

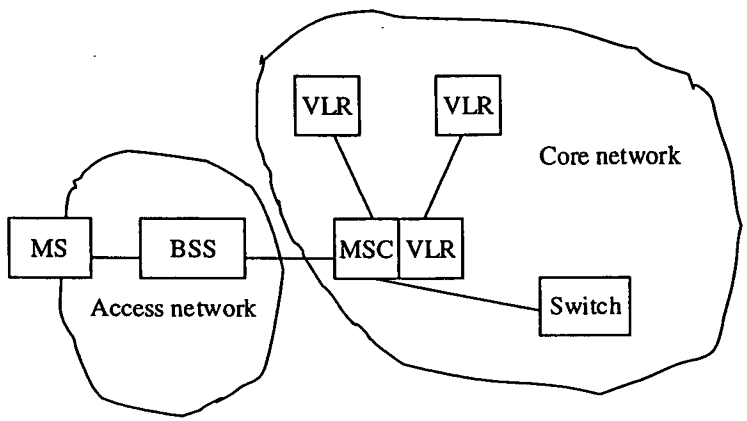


Figure 12.6 Functional network architecture with GSM entities used as an example.

12.3.1 Generic Radio Access Network Concept

In ETSI, the Generic Radio Access Network (GRAN) has been defined according to principles identified in the global mobile multimedia (GMM) report [6]. The basic idea of the GRAN is to separate the development of core and access networks. The benefit is that it leaves freedom in the future for new types of products, which utilize different core network solutions. Partly, the reasons to introduce the concept of GRAN have been political: to give freedom for standardization between rival camps. The interested reader can refer to [7] for different views about the GRAN concept.

The main focus of UMTS standardization will be in the radio access system, which will be connected to several core networks through interworking units (IWU), as depicted in Figure 12.7. The GSM BSS (Base Station Subsystem) is also part of GSM/UMTS system concept. The GRAN concept leads to an open, multivendor system that has proven to be attractive to operators. UMTS air interface will be an open and well-specified standard (U_u interface). Furthermore, the I_u interface specifies the interface between GRAN and core network, while the C_u interface specifies the interface between terminal equipment and UMTS subscriber identity module (USIM).

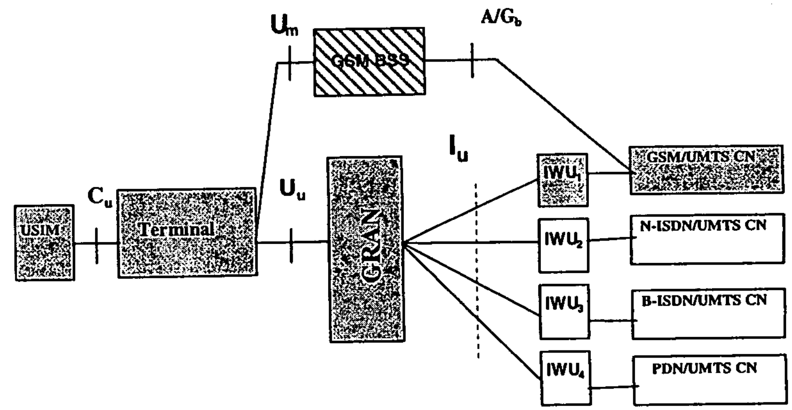


Figure 12.7 GRAN concept.

The internal structure of the GRAN comprises all the radio related functions. These include:

- All functions located in the MS terminating the lower air interface layers and controlling the channel management over the air interface from MS side;
- All functions located in the fixed part of the GRAN, called BSS, including all radio resource management functions.

In order to maintain the separation between radio dependent and independent parts, access and core networks, handover should be limited to BSS. This requires connections between BSSs.

12.3.2 ITU Family of Systems

The original idea of IMT-2000 was to create the specifications for a single global system. It has gradually been realized that changing overnight all networks to a new system would not work in practice due to commercial interests. This led to the development of the "family of systems" concept. The system belonging to ITU family of systems should be capable of supporting the IMT-2000 requirements and capabilities. Thus, the IMT-2000 target system has been replaced by a virtual IMT-2000 reference network that is primarily used for specifying the necessary interfaces for the IMT-2000 family members. The main candidates for IMT-2000 family members are GSM and IS-41 based core networks, which are evolving towards third generation.

The work to specify IMT-2000 family concept is currently underway. The ITU-T recommendation Q.FIN contains the definition of the family concept [8]. The main value of that work will most likely be the achieving of a consensus about the main third generation requirements, called capability sets, and, if desired, specification of roaming capabilities between IS-41 and GSM based IMT-2000 networks.

12.4 NETWORK TECHNOLOGIES

This section reviews some technologies that will be used in third generation networks, namely intelligent networks, asynchronous transfer mode, the signaling system 7, and mobile application part, and Mobile Internet protocol.

12.4.1 Intelligent Networks

The IN concept was developed to enable a fast deployment of new services in all telecommunication networks. The objective of IN is to allow the inclusion of additional capabilities to facilitate service provision, independent of the service/network implementation in a multivendor environment [9]. IN distinguishes between service switching, service control, and service data functions. The service control function (SCF) is the functional entity in the IN that contains the service logic for implementing a particular service. The service switching function (SWF) is the functional entity that interfaces the switch to the service control function. IN services are defined as sets of capabilities in ITU-T recommendations. Currently, Capability Set 1 (CS1) is stable.

An example of IN service is a simple number translation, like the number series "#61" would mean that your calls should be forwarded to a given number. Then, if the service provider would like to change what that number string means, the change could be easily programmed to the SCF without changing the basic telecommunication software in the switch.

12.4.2 Asynchronous Transfer Mode

Today's transmission protocols of many telecommunication networks are based on pulse code modulation (PCM), and mobile radio systems are no exception. The switching is also based on the switching of 64- or 56-Kbps PCM connections. ATM technology has received lot of attention during recent years as the next major transport technology. ATM has also been proposed for wireless applications (i.e., ATM cells are transmitted over the air interface). However, here we assume that it is only used as a transmission protocol in the infrastructure side.

ATM provides not only the multiplexing gains of packet switching, but also the guaranteed delay characteristics of circuit switching. The fundamental strategy behind ATM is to split the information into small fixed size units that are easy to handle. The fixed size of the cell allows efficient switching. ATM networks are high-speed switching systems offering large bit pipes, which allow *statistical multiplexing* (i.e., multiplexing of many connections with variable rate characteristics), which altogether reduces the overall bandwidth requirements. Since ATM is based on the transmission of fixed size cells, it can be easily evolved for future services.

The basic unit in an ATM is a cell of 53 bytes, illustrated in Figure 12.8. It consists of header (5 bytes) and information payload fields (48 bytes). The header consists of following fields:

- Generic flow control (GFC);
- Virtual path identified (VPI);
- Virtual channel identifier (VCI).

The VPI and VCI are used to identify the virtual path and virtual channels identified with that path to route the ATM cells from the source node to the destination node.

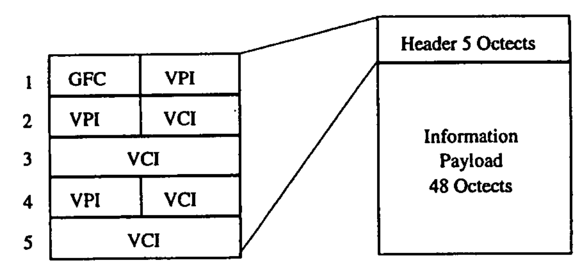


Figure 12.8 ATM cell format.

The ATM and the ATM adaptation layer (AAL) form the data link layer. AAL converts the arbitrarily formatted information supplied by the user into ATM cells. Various forms of AAL protocols are necessary to handle the different types of traffic.

- *AAL0* provides direct access to the ATM layer.
- *AAL1* assumes constant bit rate traffic, which is intolerant of missequenced information and variation in delay. It offers the following functions: segmentation and reassembly (SAR), handling of delay variation handling, handling of lost and misinserted cells, source clock recovery, monitoring for bit errors, and handling those errors.
- *AAL2* is used, for example, for voice and video. It assumes that traffic is bursty and intolerant of missequencing and that a time stamp is needed for packet reassembling. It offers the following functions: multiplexing, SAR, handing delay variation, handling cell lost/error, and source clock recovery. AAL2 will be used for compressing data in third generation mobile radio systems in the network infrastructure.
- *AAL3/4* and *AAL5* are geared to traffic that has bursty characteristics with variable frame length. Furthermore, delay is not critical and packets can be resequenced based on sequence numbers. AAL5 is expected to supersede AAL3/4 since it has lower overhead and TCP/IP acknowledgements fit into one cell in AAL5 instead of two cells in AAL3/4.

12.4.3 Signaling System 7

Both IS-41 and GSM make use of protocols specified for SS7, which is a signaling system designed for the transfer of control information between network elements. The layered structure of SS7 is shown in Figure 12.9. SS7 layers are called parts. The ISDN user part (ISUP) of SS7 contains messages carried from ISDN standard devices. For example, call-related signaling of GSM makes use of ISUP when connecting to external networks. The GSM and IS-41 MAP protocols use the transaction capabilities part (TCAP) for network control. Furthermore, the signaling connection control part (SCCP) and message transfer part (MTP) are used in the A-interface of GSM and in the IS-634 interface, which specifies the corresponding A-interface for the North American digital standards, especially for IS-95/IS-41. A tutorial of SS7 can be found in [10], and use of SS7 in IS-41-based systems is discussed in [11].

12.4.4 Mobile Application Part

A unique feature of mobile systems is the support of roaming (i.e., seamless provision of services for users who are subscribed to one operator's network while they are within the coverage of some other operator's network). The functions required for roaming are authentication of the subscriber, transferring subscriber data to the visited network, and mechanism for routing connections towards the subscribers. The functions to implement roaming are part of the mobility management. For this purpose, both GSM and IS-41 systems have defined mobile application part (MAP). MAP defines the application protocols between switches and databases (e.g., MSC, VLR, HLR) for supporting call management, supplementary service management, short message transfer, location management, security management, radio resource management, and mobile equipment

management. In principle, the MAP of IS-41 and GSM perform similar functions. However, their implementations and the way that protocols are specified are different.

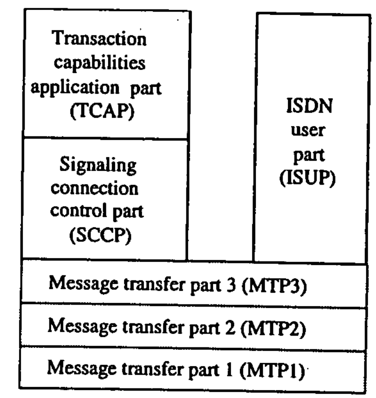


Figure 12.9 SS7 protocols.

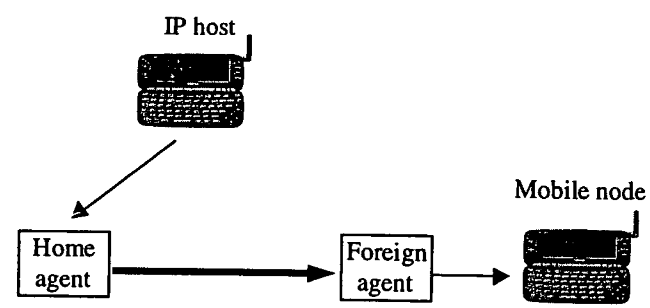


Figure 12.10. Mobile IP architecture.

12.4.5 Mobile Internet Protocol

In general, it can be stated that Mobile Internet protocol (IP) is for data networks what MAP is for telecom networks. Internet is built using software that relies on IP. Mobile IP is a modification to IP, which makes possible for a mobile node to visit other IP networks while still being reachable by its home IP address. Mobile IP architecture is based on the *home/foreign agent* concept depicted in Figure 12.10. The datagrams are routed to the home network using normal IP routing. If the mobile node is in a foreign network, the datagrams are encapsulated and tunneled towards the so-called *care-of*

address assigned to the mobile node when away from the home network. For more details on Mobile IP, refer to [12,13].

12.5 EVOLUTION OF SECOND GENERATION NETWORKS

In the beginning of third generation system standardization, it was not clear that the evolution from second generation networks would be the right way to proceed nor even what was meant by the evolution of second generation networks. The original idea was to start from scratch and to define a completely new core network and protocols for mobility management and call control. However, during recent years it has become more evident that third generation core networks will be based on the evolution of second generation networks. This is due to the desire to utilize existing network investments, and also because many of the original goals defined for third generation networks have already been implemented in second generation networks. However, one still has to consider a balance between evolution, which is restricted by the existing implementations of the standard, and revolution, which can be done without limitations and thus results in a more efficient system.

In order to develop the technical solutions for third generation networks, we need to answer two questions: in what areas is evolution needed and how can this evolution be implemented. Therefore, we need to identify the capabilities that differentiate third generation from second generation networks. Once we have identified the new capabilities that need to be supported, we can analyze what changes and new functionalities are required in the existing systems and standards. Next, the distribution of these functionalities into protocols and network elements, along with possible new protocols, network elements, and interfaces, needs to be identified. We have to separate those aspects that impact the standard and those that impact the way networks are implemented. For example, manufacturers of wireless equipment usually design the detailed algorithms for radio resource management.

12.5.1 New Capabilities of Third Generation Systems

In this section, we discuss what the impact of new services is on the required system capabilities.

12.5.1.1 Multimedia

Third generation networks will support multimedia services, meaning sufficient bandwidth and bearer flexibility [14]. The first requirement, sufficient bandwidth, relates to transport technology. Since some third generation services are asymmetric, bursty, and demand high bandwidth, transmission capabilities need to be developed to match these requirements. Higher air interface rates together with variable bit rates will mean that new, more flexible transport and switching technology is required. The support of multimedia service calls for separation of the call control from the connection and bearer control. A call/session might use various connections at any one

particular instant. Thus, it should be possible to add and remove bearers during such a call. Separation of call and connection/bearer control means that many connections such as speech, video, and data could be associated with one single call and these could be handed over separately [14]. Thus, handover algorithms need to support multiple simultaneous services and different handover schemes for them.

One of the main applications for third generation networks is assumed to be the transmission of Internet services. Currently, the IP makes no assumptions about the underlying protocol stacks and offers an unreliable, connectionless network-layer service that is subject to packet loss, reordering, and packet duplication [15]. Therefore, the IP-delivery model is referred to as a best-effort service. For non real-time services, this kind of transmission model is not a main problem. However, for multimedia services, it is not adequate, and enhanced QoS classes are required. The resource reservation protocol (RSVP) is specified to provide means to reserve network resources along the data path, and to ensure end-to-end QoS for the selected application. It could be a solution for this problem. Other protocols are also considered for this purpose [15].

12.5.1.2 Bearer Service Classification

In order to better understand what kind of new bearer services are required, we introduce the ATM service classification. The third generation service classes are not necessarily exactly the same, but, since ATM was designed to carry all kinds of traffic, it serves as a good starting point for wireless networks.

Constant Bit Rate Service (CBR): CBR may be used for any transparent data transfer. Resources are allocated on a peak bit rate basis.

Unspecified Bit Rate Service (UBR): UBR bearers use free bandwidth when available. If no resources are available, the information frames are queued.

Available Bit Rate Service (ABR): Resources for sending at the specified minimum bit rate are allocated to the user. Higher bit rates may be used on best effort basis, free bandwidth is used when available.

Variable Bit Rate Service (VBR): This service provides a variable rate characteristic based on statistical traffic management. Average throughput (sustainable bit rate) and frame loss ratio are two key parameters. A set of traffic parameters specify the source traffic characteristics. This service can be either real-time (with delay bound) or non real-time (without delay bound).

Currently, second generation networks support CBR and UBR (with and without priority) services. GSM HSCSD also supports ABR services. VBR service would require enhancements to second generation core networks.

12.5.1.3 *Service Development*

One of the major differences between second and third generation networks is that instead of standardized services, standardized bearer capabilities for supporting services are provided. This means that rather than standardizing teleservices, bearer services providing suitable "bit-pipe" for any kind of services will be specified. Nevertheless, the most common services like speech would still have to be standardized to ensure seamless roaming and spectrum efficiency.

A very important protocol for mobile systems is the mobility management, which facilitates seamless roaming. This requires certain agreements and also that networks support the same services. On the other hand, due to hard competition, operators would like to distinguish themselves from each other. This has led to the emergence of operator-specific services provided by INs. With IN the operator-specific services can be provided in a flexible way

A further extension of the IN concept is the virtual home environment (VHE) [16]. It is defined as a system concept for personalized service portability across network boundaries. The VHE concept will ensure uniform appearance and presentation of services and features. Thus, not only can the user access services in a similar way as in the home network, but, for example, response messages or error messages are also presented in a similar way.

12.6 GSM EVOLUTION TOWARDS UMTS

In this section, we analyze the existing architecture, interfaces, and protocols of the GSM and GPRS networks. Based on this, we introduce a possible evolution scenario and suggest some enhancements required to allow the evolution of GSM and GPRS towards the third generation.

12.6.1 Reference Architectures

For GSM, we need to consider two reference architectures, one for the basic GSM and one for GPRS to identify the interfaces subject to evolution. These are illustrated in Figures 12.11 and 12.12, respectively.

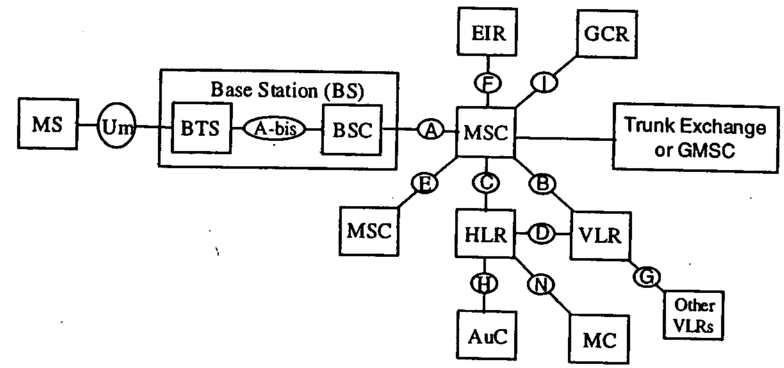


Figure 12.11 Basic GSM reference architecture.

In addition to the network elements specified in Section 12.2.3, we can identify the following GSM-specific network elements.

Subscriber Identity Module (SIM): This is part of the mobile station and is used to store all subscriber related information. It is either a smart card (the size of a credit card) or a so-called "plug-in" SIM, which was designed to facilitate an easier implementation of small terminals.

Group Call Register (GCR): The GCR is the management database for the voice group or broadcast calls in the area controlled by the associated MSC(s). Whenever the MSC needs data related to a requested voice group or broadcast call, it interrogates the GCR to obtain the respective voice group or broadcast call attributes. More information is provided in TS GSM 03.68 and 03.69.

Shared Interworking Function (SIWF): The SIWF is a network function that provides interworking for circuit data/fax calls. SIWF consists of a SIWF controller (SIWFC) functionality located in MSCs and SIWF server(s) (SIWFS) located in the PLMN. A SIWFS can be accessed by several other network nodes (e.g., any MSC in the same PLMN). More information is provided in TS GSM 03.54.

Gateway MSC (GMSC): If a network delivering a call to the PLMN cannot interrogate the HLR, the call is routed to an MSC. This MSC will interrogate the appropriate HLR and then route the call to the MSC where the mobile station is located. The MSC that performs the routing function to the actual location of the MS is called the gateway MSC (GMSC). For more information, refer to TS GSM 03.04.

SMS Gateway MSC (SMS-GMSC): The SMS-GMSC acts as an interface between a short message service center and the PLMN, to allow short messages to be delivered from the service center (SC) to mobile stations.

SMS Interworking MSC: The SMS interworking MSC acts as an interface between the PLMN and a short message service center to allow short messages to be submitted from mobile stations to the SC.

Interworking Function (IWF): The IWF is a functional entity associated with the MSC. The IWF provides the functions necessary to allow interworking between a PLMN and the fixed networks (ISDN, PSTN, and PDNs). The functions of the IWF depend on the services and the type of fixed network. The IWF is required to convert the protocols used in the PLMN to those used in the appropriate fixed network. The IWF may have no functionality where the service implementation in the PLMN is directly compatible with that at the fixed network. The interworking functions are described in GSM Technical Specifications GSM 09.04 – 09.

Interfaces: In the basic GSM configuration presented in Figure 12.10, all the functions are considered implemented in different equipments. Therefore, all interfaces within PLMN are external. Interfaces A and Abis are defined in the GSM 08-series of technical specifications. Interfaces B, C, D, E, F, and G need the support of the SS7 MAP to exchange the data necessary to provide the mobile service. No protocols for the H-interface and for the I-interface are standardized. We now describe the main interfaces defined between the different GSM network elements:

- The interface between BSC and the MSC (A-interface) is specified in the following GSM specifications: GSM 08.01, 08.02, 08.04, 08.06, 08.08, and 08.20.
- The interface between the BTS and the BSC (the Abis-interface) is specified in the following GSM specifications: GSM 08.51, 08.52, 08.54, 08.56, 08.58, 08.59, and 08.60.
- The interface between the MSC and the HLR (C-interface): signaling in this interface uses the MAP protocols specified in GSM 09.02.
- The interface between the MS and the BTS (air interface, Um-interface) is specified in the GSM 04 and 05 specification series. The GSM 04 series covers the general principles and higher protocol layers, and the GSM 05 series contains the physical layer specifications.
- The interface between the MSC and its associated VLR (B-interface) is internal to the MSC/VLR, and signaling on it is not standardized.
- The interface between the HLR and the MSC (C-interface): the GMSC must interrogate the HLR of the required subscriber to obtain routing information for a call or a short message directed to that subscriber. Signaling in this interface uses the MAP protocols specified in GSM 09.02.
- The interface between the HLR and the VLR (D-interface) is used to exchange data related to the location of the mobile station and to the management of the subscriber. Signaling in this interface uses the MAP protocols specified in GSM 09.02.

- The interface between MSCs (E-interface) is used for the exchange of data for handover between two MSCs. This interface is also used to forward short messages. Signaling on this interface uses the MAP specified in GSM 09.02.
- The interface between VLRs (G-interface): when an MS initiates a location updating using TMSI, the VLR can fetch the IMSI and authentication set from the previous VLR. Signaling on this interface uses the MAP specified in GSM 09.02.
- The interface between the HLR and the gsmSCF (J-interface) is used by the gsmSCF to request information from the HLR. The support of the gsmSCF-HLR interface is a network operator option. As a network operator option, the HLR may refuse to provide the information requested by the gsmSCF.

GPRS architecture: GPRS is logically implemented on the GSM structure through the addition of two network nodes, the serving GPRS support node (SGSN) and the gateway GPRS support node (GGSN) [17]. Therefore, several new interfaces have been defined. Figure 12.12 depicts these, together with the GSM interfaces explained previously.

- The interface between the SGSN and the BSC (Gb-interface).
- The interface between GGSN and HLR (Gc-interface) is optional.
- The interface between SMS-GMSC and SGSN, and between SMS-IWMSC and SGSN (Gd-interface) enables GPRS MSs to send and receive short messages over GPRS radio channels.
- The interface between two GSNs within the same PLMN (Gn-interface).
- The interface between two GSNs in different PLMNs (Gp-interface) has the same functionality as Gn-interface plus a security function required for inter-PLMN communication.
- The interface between SGSN and HLR (Gr-interface).
- The interface between SGSN and MSC (Gs-interface).
- The interface between GPRS and fixed network (Gi-interface) is actually called a reference point since it is not fully specified. This is described in GSM 09.61.

For more details on GPRS, refer to [18–20].

12.6.2 Protocol Stacks

Figure 12.13 illustrates the GSM protocol stack. The layer 3 protocols – radio resource management (RRM), mobility management (MM), and call control (CC) – are specified in GSM 04.08. LAPDm and LAPD are link layer protocols. SCCP′ is the GSM specific implementation of the original SS7 SCCP protocol [4]. The A-interface is used for messages between BSC and MSC as well as for messages to and from the mobile station. Therefore, the BSS application part (BSSAP) specified in GSM08.08 is split into two subapplication parts, the BSS management application part (BSSMAP) and the

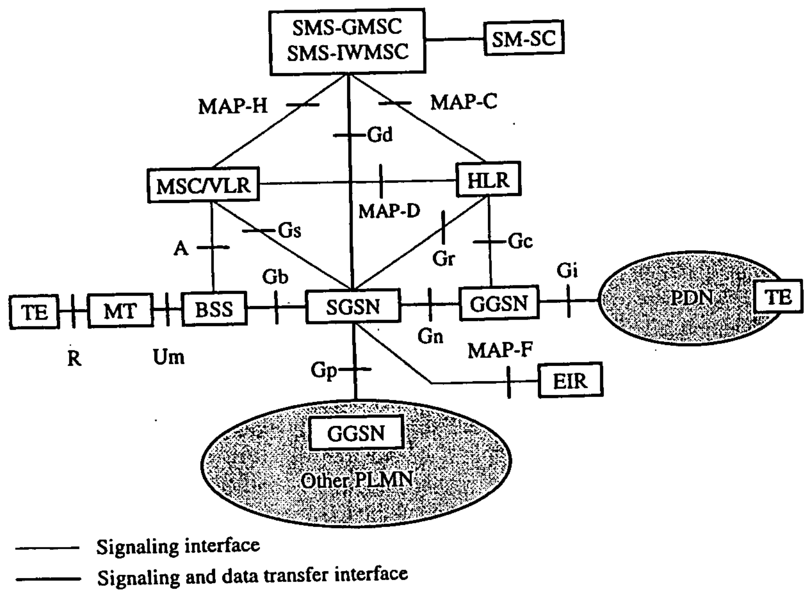


Figure 12.12 GSM/GPRS reference architecture. (*Source*: [17], reproduced with permission from ETSI.)

direct transfer application part (DTAP). DTAP is used to transfer the signaling for call and mobility management (CC and MM protocols) transparently to the mobile station. The BSSMAP supports all the procedures between the MSC and the BSS that require interpretation and processing of information related to single calls, and resource management. Both these protocols make use of the SCCP and MTP protocols, which are part of the SS7 protocols. The transaction capabilities application part (TCAP) is used by the MAP protocol for network control. Not shown in Figure 12.13 is the MAP protocol. GSM 09.02 specifies the application protocols between exchanges and data bases (MSCs, GMSCs, VLRs, HLRs, EIRs, and SGSNs) for supporting call management, supplementary service management, short message transfer, location management, security management, radio resource management, and mobile equipment management. It also specifies the applicability of SCCP and TCAP protocols to support these exchanges.

12.6.2.1 GPRS Protocol Stacks

The GPRS protocol architecture makes a distinction between the transmission and signaling planes. The transmission plane consists of a layered protocol structure providing user information transfer, along with associated information transfer control procedures (e.g., flow control, error detection, error correction, and error recovery).

Figure 12.14 shows the protocol stack for the GPRS transmission/user plane, while Figure 12.15 presents the GPRS signaling/control plane.

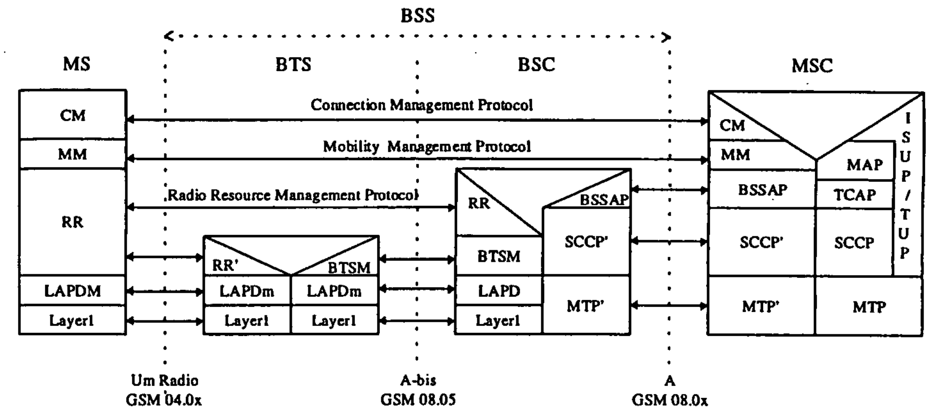


Figure 12.13 GSM protocol stack.

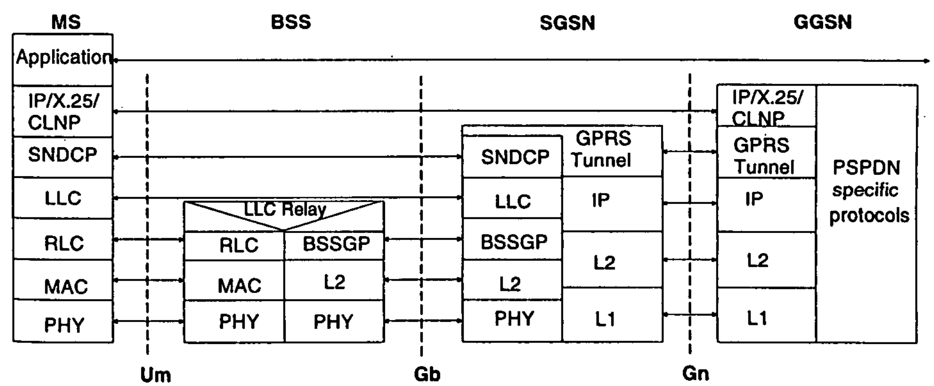


Figure 12.14 GPRS transmission/user plane. (*Source*: [17], reprinted with permission of ETSI.)

The GPRS tunneling protocol (GTP) transfers data between GPRS support nodes. It encapsulates the protocol data units (PDUs) of packet data protocols such as X.25 or IP. Furthermore, GTP provides a mechanism for flow control between GSNs, if required. The GTP protocol is defined in GSM 09.60. Below the GTP protocol are TCP and UDP protocols. TCP carries GTP PDUs for protocols that need a reliable data link (e.g., X.25) and UDP for protocols that do not need a reliable data link (e.g., IP). IP

version 4 is used as a GPRS backbone network-layer protocol providing routing functions for user data and control signaling. The IP version 6 will replace version 4 in the future.

Between the SGSN and the MS, packet data protocol (PDP) PDUs are transferred with the subnetwork dependent convergence protocol (SNDCP) that maps network-level characteristics into the characteristics of the underlying logical link, and provides multiplexing of multiple layer 3 messages into a single virtual logical link connection. In addition, ciphering, segmentation, and compression are also performed by the SNDCP protocol. Logical link control (LLC) provides reliable logical link. The LLC relay function relays LLC PDUs between Um and Gp interfaces. The base station system GPRS protocol (BSSGP) conveys routing and QoS-related information between BSS and SGSN. It is specified in GSM 08.64.

In the radio interface GSM RF, the GSM physical layer specified in GSM 05 series performs modulation, demodulation, encoding, and decoding of data.

The data link layer has two sublayers: radio link control (RLC) and medium access control (MAC). RLC provides a radio solution-dependent reliable link. The MAC function controls the access signaling (request and grant) procedures for the radio channel, as well as the mapping of LLC frames into the GSM physical channel. The RLC and MAC functions are described in GSM03.64.

The signaling plane consists of protocols for control and support of the transmission plane functions such as attaching to and detaching from the GPRS network, activation of a PDP (e.g., IP or X.25) address, controlling the routing path to support user mobility, controlling the assignment of network resources, and providing supplementary services.

The previously described physical layer, RLC/MAC and LLC protocols are also used for signaling transmission between the MS, BSS, and SGSN. The layer 3 mobility management (L3MM) protocol supports mobility management functionality such as GPRS attach, GPRS detach, security, routing update, location update, PDP context activation, and PDP context deactivation between MS and SGSN.

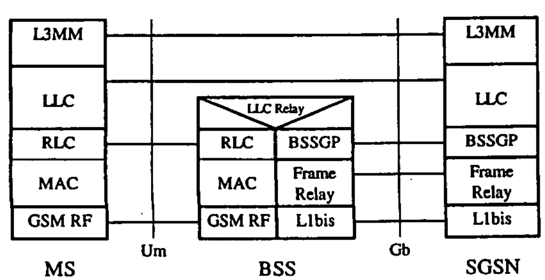


Figure 12.15 GPRS signaling/control plane between MS and SGSN. (*Source*: [17], reprinted with permission of ETSI.)

Signaling between the SGSN and HLR is performed using the same protocols, TCAP, SCCP and MTP, as for the non-GPRS GSM PLMNs. The MAP protocol supports signaling exchanges between SGSN and HLR with enhancements for GPRS.

12.6.3 Evolution of the GSM Architecture

In this section, we discuss the expected changes to GSM architecture and protocols due to the evolution towards third generation. These changes are currently discussed and developed in the standardization, and the reader is referred to the standardization documents in ETSI SMG3 AND SMG12 (see Chapter 14) for the latest information.

As shown in Figure 12.16, the first implementations of GRAN will be based on the integration of URAN and GSM/UMTS core network, which has been evolved from the GSM core network by integrating new third generation capabilities. The evolved GSM network elements are referred to as 3G MSC and 3G SGSN.

GRAN interfaces with GSM/UMTS core network via Iu-interface corresponding to GSM A-interface and GPRS Gb-interface [21]. As can be seen, radio access is isolated from the core network, and the goal is that the GSM/UMTS core network would have the flexibility to support any radio access scheme. Circuit switched services are routed via the GSM MSC, and the packet switched services via the GPRS part of the GSM/UMTS core network.

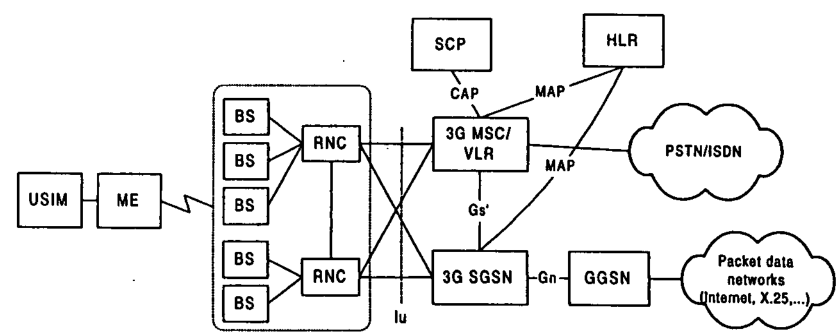


Figure 12.16 One possible scenario where GRAN is connected to an evolutionary GSM core network; 3G MSC/VLR and 3G SGSN provide the Iu interface.

It is not yet clear how tightly the Iu-interface will combine interfaces A and Gb. However, it seems desirable that they could be more integrated than today. This can be motivated from the service perspective. Today's clear separation of speech and circuit switched and packet data might not be true in the future. Developments such as RSVP make it possible to provide guaranteed QoS for packet networks, and thus, if UMTS/GSM is used to connect a user to packet access networks, similar capabilities need to be provided.

The question is also whether SGSN and MSC will be combined or will they remain separate entities. In this case, we would have within the same call connections to both the packet data side and the circuit data side. This would mean that MSC and SGSN would have to request abstract communication resources (i.e., in principle, QoS requirements only, not actual radio resources), and BSC would allocate radio resources for the communication according to current needs. As an example, a packet-based use of radio resources for IP traffic could be upgraded any time to circuit-based (without informing SGSN about it) if the MSC side wants to establish a call. The BSC would thus find the minimum radio resources that fulfil the sum of all requirements of MSC and SGSN services (the highest QoS requirement obviously influences the selection the most). Of course, if SGSN also supports other than best effort services, then it could also request higher priority services.

The clearer separation of the radio access and core networks in UMTS, when compared to GSM, results in some questions that need to be solved before paging procedures, and consequently the required signaling can be determined. For example, in what level does the core network need to know about the MS location inside GRAN? If the level is high, then this results in high traffic between the core network and GRAN when the mobile station is moving [5].

12.6.3.1 Handover

It is assumed that the handover decision is always made inside the GRAN. Thus, for inter-GRAN handover, functions to set up a path within the core-network, or some other arrangements, are required. Depending on the inter-GRAN handover type, different changes for GSM are required. The backward handover, where the handover signaling is performed through the old base station, is very similar to the current GSM handover. In the forward handover, the mobile station initiates the handover through the new base station. The need for this type of handover was discussed in Chapter 11: when trying to avoid corner effect, in which the connection to the old base station is lost, a very fast handover is required to prevent the blocking of the existing users in another cell. Forward handover requires a large number of changes in GSM.

To speed up handovers between two RNCs, direct interconnections similar to IS-634 (see Section 12.7.2) have been proposed. For this, a new interface between RNCs, as shown in Figure 12.16, is required. For services which have long packets or frequently transmit small packets, there is a physical connection established. Thus, a handover procedure is also required for packet service, except for short packets, which are transmitted in the random access message. If handover is performed between two BSC belonging to different GPRS support nodes, then there has to be a way to establish soft handover between these RNCs.

12.6.3.2 Transmission Infrastructure

The transmission infrastructure has to meet the new requirements imposed by wideband services. Since the data services are bursty and often asymmetric, the transmission solution has to be able to efficiently multiplex different types of information. ATM will provide efficient support for transmission of bursty wideband services. However, since

ATM was originally designed for very high-speed transmission in the fixed network, some modifications may be needed to accommodate cellular-specific infrastructure requirements.

12.6.4 Protocol Aspects

Figure 12.17 shows a possible protocol stack for UMTS, and Figure 12.18 shows the GRAN connected to a packet data network using the evolved GPRS protocols. Obviously, since there will be a new air interface, wideband CDMA, the radio-dependent protocols such as layers 1 and 2 will be completely replaced. The upper layer protocols, on the other hand, can be reused with some adaptation. For example, CC and MM protocols can be adapted, and thus, GSM 04.08 could form the basis for further development of UMTS CC and MM protocols denoted as MM' and CC' in Figure 12.17. The RR messages as such are not radio dependent and can be at least partly reused. The GPRS LLC and GSM RLP could be used as such for the wideband CDMA air interface. If ATM is used as the basic transport technique in the network side, the LAPD protocol will be replaced by AAL2/5, service specific convergence sublayer (SSCF), and service-specific connection oriented (SSCOP) protocols.

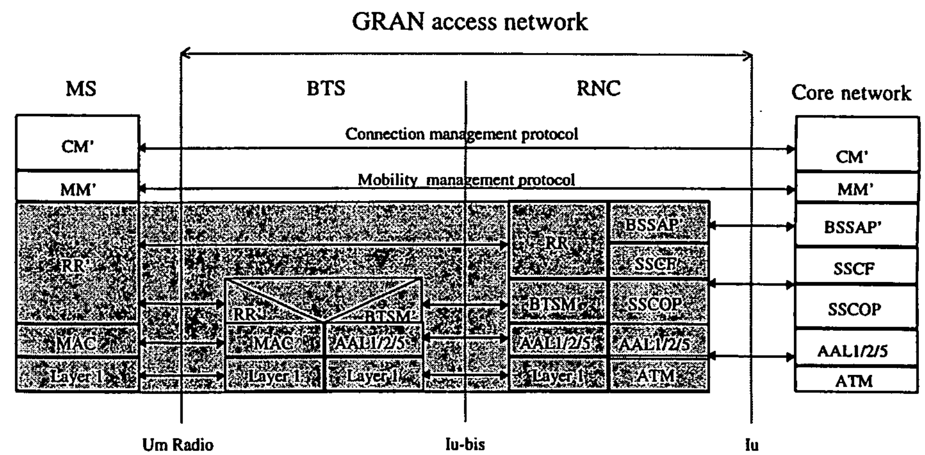


Figure 12.17 General protocol stacks of GRAN using GSM as a core network. The GRAN protocols are shaded.

12.6.4.1 Evolution of GSM BSSMAP and BSSGb Protocols

We can distinguish two different categories of functions over the GSM A-interface. The first category covers the functions to the BSS, while the second covers the functions to

the MS. The functions related to the first category are used mainly to manage radio resources and are as follows:

- Bearer setup;
- Bearer release;
- Resource check;
- Paging;
- Ciphering;
- Handover.

BSSMAP (part of the BSSAP protocol) and BSSGP protocols specify these functions. The new BSSAP and BSSGP protocols are denoted by BSSAP' and BSSGP' in Figures 12.17 and 12.18. The information transmitted in messages supporting the above functions is specified by data elements. These data elements need to be updated. For example, bearer control messages need to be specified to include new channel types. There will then be new bit rates and new coding schemes. Furthermore, not all coding schemes are known today. The need for some generic bearer service classes exists, since wideband CDMA could support an almost infinite number of data rates due to the very high granularity. As there will be more mobile station classes due to the increased number of service possibilities, the GSM classmark information carrying the properties of the mobile stations should also be updated.

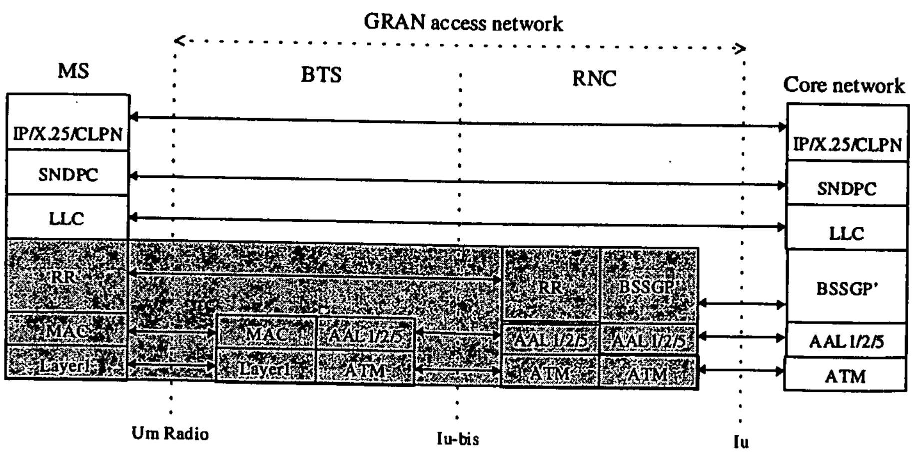


Figure 12.18 Transmission/user plane for the packet data interconnection.

Wideband CDMA GRAN will support QoS negotiation. Furthermore, as was discussed previously, it is expected that packet data networks will also evolve in this

aspect, and thus QoS negotiation support from GPRS backbone network will be needed as well. Thus, the main change is that the GPRS BSSGB protocol should support some extra control messages for QoS negotiation purposes.

12.6.4.2 CAMEL

The purpose of Customized Applications for Mobile Network Enhanced Logic (CAMEL) is to provide a mechanism for supporting services consistently and independently of the subscribers' location [22]. It is based on IN principles and can be used to provide operator-specific services for roaming users. CAMEL relies on triggers within the supporting network infrastructure to suspend call processing and communicate with a remote computing platform before proceeding to handle the service functions. CAMEL uses the GSM MAP protocol to transfer information between appropriate network elements. The service switching function (SSF) in the visited network will have an interface directly to the service control function (SCF) in the home GSM network.

Non call-related events such as call independent supplementary service procedures, transfer of short messages, or mobility management procedures are within CAMEL's scope [22]. Thus, there is still need for further evolution towards the full implementation of the VHE concept, which would manage full transparency of offered services regardless of the network to which a user is currently attached.

12.7 EVOLUTION OF IS-41/ IS-95

In this section, we analyze the existing architecture, interfaces, and protocols of the IS-41/IS-95 networks. Based on this, we introduce a possible evolution scenario and suggest some of the required enhancements to allow the evolution of IS-95/IS-41 towards the third generation. An overview of the different standards required in the implementation of IS-41/IS-95-based systems is given in the next section. For a detailed treatment of the IS-41 standard, refer to [23].

12.7.1 Overview of IS-41/IS-95 Standards

When considering a complete IS-95 based network implementation, we need to consult several standards. The main standards are:

- EIA/TIA/IS-95B, Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System;
- EIA/TIA/IS-41-C, Cellular Radiotelecommunications Intersystem Operations;
- TIA/EIA/IS-707, Data Service Options for Wideband Spread Spectrum Systems;
- TIA/EIA/IS-658, Data Services Interworking Function Interface for Wideband Spread Spectrum Systems;

- TIA/EIA/IS-634, MSC-BS Interface, Rev A.

12.7.2 Reference Architecture

Figure 12.19 shows a partial view of the reference architecture for the IS-41 core network. The message center (MC) is an entity that stores and forwards short messages. It may also provide supplementary services for SMS.

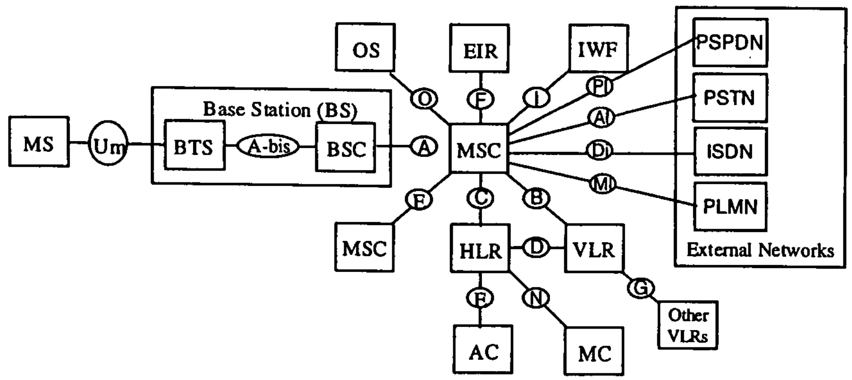


Figure 12.19 Partial view of IS-41 reference architecture.

The following main interfaces are defined between the network elements:

- The interface between the BSC and the MSC (A-interface) is specified in the IS-634 standard.
- The interface between the BTS and the BSC (the Abis-interface) is not specified.
- The interface between the MSC and the HLR (C-interface): signaling is specified in the IS-41 standard.
- The interfaces between the MS and the BTS (air interface, Um-interface) are specified in the IS-95, IS-97, and IS-98 standards.
- The interface between the MSC and its associated VLR (B-interface) is not specified.
- The interface between the HLR and the MSC (C-interface): signaling is specified in the IS-41 standard.
- The interface between the HLR and the VLR (D-interface) is used to exchange data related to the location of the mobile station and to the management of the subscriber. Signaling in this interface is specified in the IS-41 standard.
- The interface between the HLR and the VLR (D-interface): signaling is specified in the IS-41 standard.

- The interface between IWF and MSC (L-interface) is specified in the TIA/EIA/IS-658 standard.

12.7.2.1 Detailed A-interface Architecture

The first version (Rev 0) of the IS-634 standard defining the A-interface was published in 1995. The first implementations of IS-95-based systems were thus actually proprietary and, consequently, the standard was heavily influenced by the existing implementations. In the new revision of IS-634 Rev A, to be published in 1998, two distinct architectures have been defined: architecture A (with the selection/distribution unit (SDU) located inside of the base station) and architecture B (with the SDU located outside of the base station). Architecture A has a direct inter-BS connection option and an indirect (through MSC) BS-to-BS connection option. Architecture B has only a direct inter-BS connection.

There are six "parts" in the IS-634 Rev A. As compared to the earlier version of IS-634, a major restructuring has been performed:

- Part 0 (Base Part): high level function overview;
- Part 1 (Common Protocol): common protocol for architecture A and B;
- Part 2 (Architecture A): specific protocol for architecture A only;
- Part 3 (Architecture B): specific protocol for architecture B only;
- Part 4 (Interworking): interworking protocol between architecture A and B;
- Part 5 (Protocol Details): defines messages, information elements, and timers.

The detailed A-interfaces for architectures A and B are shown in Figures 12.20 and 12.21, respectively. The A-interface consists of A1- through A7-interfaces. A3- and A7-interfaces are both TCP/IP connections over ATM and are used for direct inter-BS connections for faster and more efficient soft handover. The A3-interface is divided into A3s (for signaling) and A3t (for traffic). Direct inter-BS connections (A3t, A3s, and A7) support faster soft handover and can add/drop multiple cells in same message.

The A6-interface is used in the indirect (through MSC) BS-to-BS connection option of architecture A (not shown). It is a 16-Kbps subrate circuit carrying user data/speech between the BS and the MSC during soft handover.

SDU is a new functional unit in IS-634. SDU contains functional entities such as traffic handler, signaling layer 2, multiplex sublayer, frame selection/distribution for soft handover, and power control. It also contains the transcoder functionality, which can be located either inside or outside of the base station.

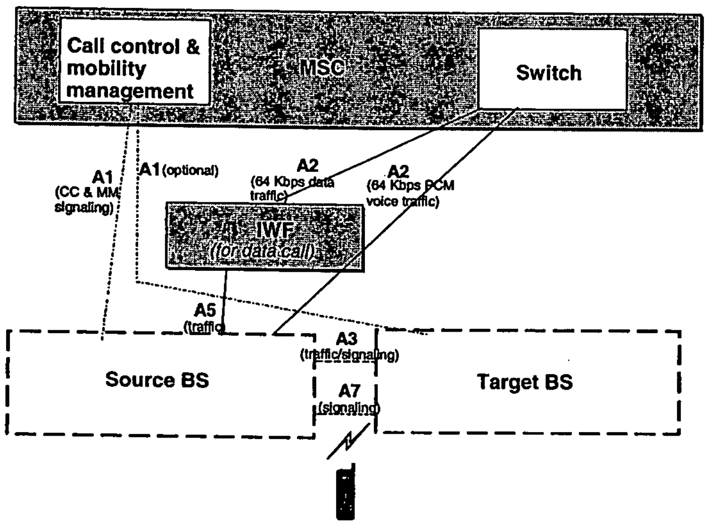


Figure 12.20 IS-634 architecture A.

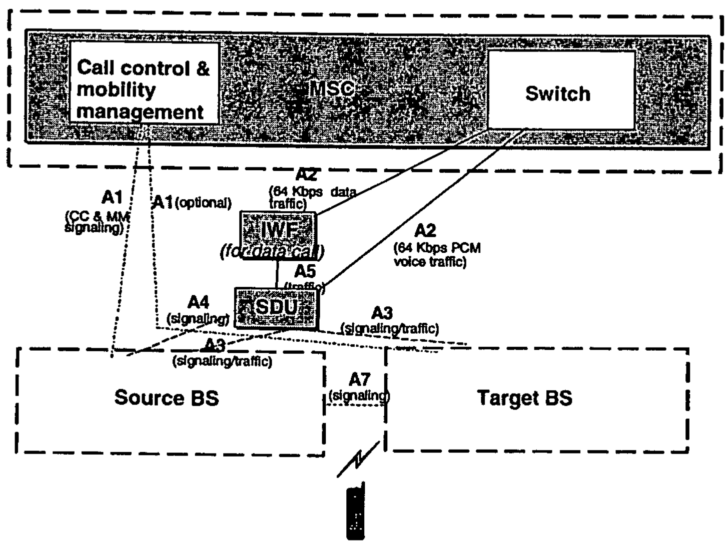


Figure 12.21 IS-634 architecture B.

12.7.3 Protocol Stacks

12.7.3.1 Air Interface

Figure 12.22 shows the IS-95 air interface protocol structure for the mobile station and the base station. Layer 1 is the physical layer of the digital radio channel, including those functions associated with the transmission of bits, such as modulation, coding, framing, and channelization via radio waves. Between layer 1 and layer 2 is a multiplex sublayer containing the multiplexing functions that allow sharing of the digital radio channel for user data and signaling processes. The multiplex sublayer provides several multiplex options. By *multiplex option* it is meant the ability of the multiplex sublayer and lower layers to be tailored to provide special capabilities. The multiplex option defines such characteristics as the frame format and the rate decision rules. The physical layer and multiplex sublayer are specified in Chapters 6 and 7 of the IS-95 specifications. 7. The IS-95 combines the operation of the network and data link layers and treats them as one layer.

Primary traffic is the main traffic stream carried between the mobile station and the base station on the traffic channel. Secondary traffic is an additional traffic stream that can be carried between the mobile station and the base station on the traffic channel. Signaling traffic means control messages that are carried between the mobile station and the base station on the traffic channel.

For user data, protocol layering above the multiplex sublayer is service option dependent and is described in standards for the service options. Service option means a service capability of the system. A *service option* may be an application such as voice, data, or facsimile.

For the signaling protocol, two higher layers are defined. Signaling protocol layer 2 is the protocol associated with the reliable delivery of layer 3 signaling messages between the base station and the mobile station, such as message retransmission and duplicate detection. Signaling layer 3 is the protocol associated with call processing, radio channel control, and mobile station control, including call setup, handover, power control, and mobile station lockout. Layer 3 signaling is mainly passed transparently to the MSC through the BS.

12.7.3.2 A-Interface

The IS-634 standard defines the MSC-BS messages, message sequencing, and mandatory timers at the base station and the mobile switching center. Figure 12.23 depicts IS-634 functions: call processing and supplementary services, radio resource management, mobility management, and transmission facilities management.

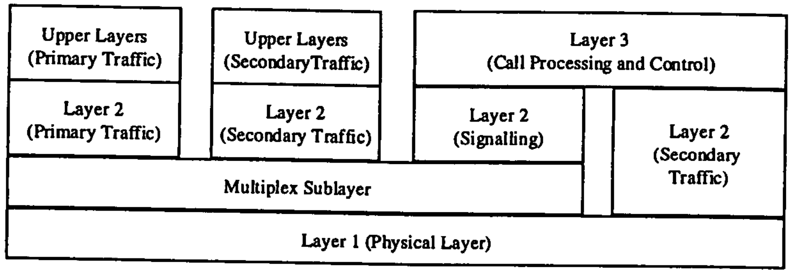


Figure 12.22 Mobile station and base station layers. (*Source*: [24], reproduced under written permission of the copyright holder (Telecommunications Industry Association). At the time of the publication, the standard which contains this figure was not finalized, please check with TIA for the correct version.)

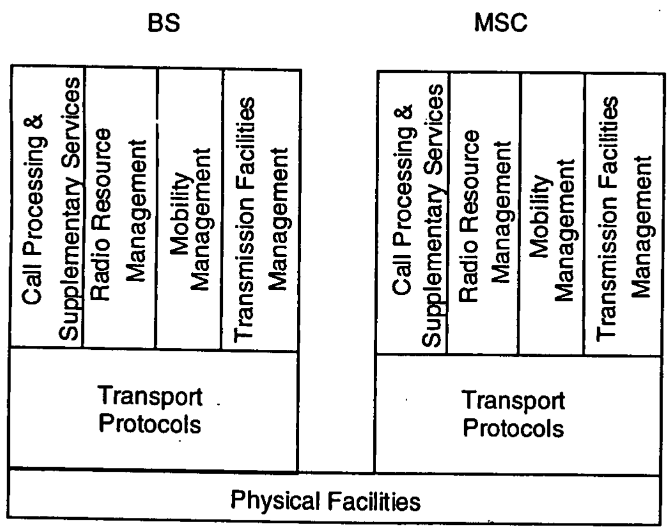


Figure 12.23 A-interface functions (*Source*: IS-634 Rev. A, reproduced under written permission of the copyright holder (Telecommunications Industry Association).)

Figure 12.24 shows the A1-interface signaling protocol reference model. Similar to GSM, IS-634 has a base station application part (BSAP), which contains BSMAP and DTAP. The BSAP corresponds to BSSAP in GSM and BSMAP to BSSMAP. Thus, DTAP is used to transfer the mobility management and call control related signaling between the MSC and the mobile station. BSMAP transfers, for example, radio resource management related signaling between BS and MSC. The transport protocols for user and signaling data are shown in Figures 12.24 and 12.25, respectively. The physical layer is based on the use of T1 digital transmission system interface (1.544 Mbps providing 24×56 or 64-Kbps channels), E1 digital transmission interface (consisting of 30×56 or 64Kbps channels), OC1 digital transmission interface (51.84 Mbps), and OC3 digital transmission interface (155.52 Mbps).

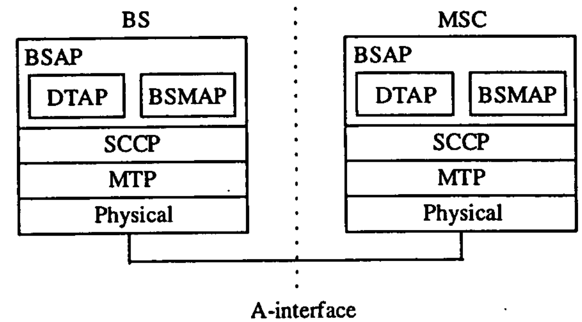


Figure 12.24 A1 interface signaling protocol reference model.

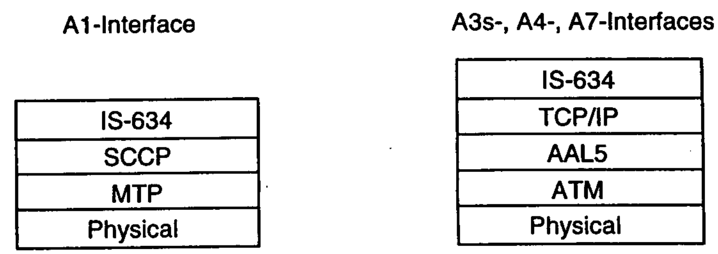


Figure 12.25 Transport protocol options for signaling connections.

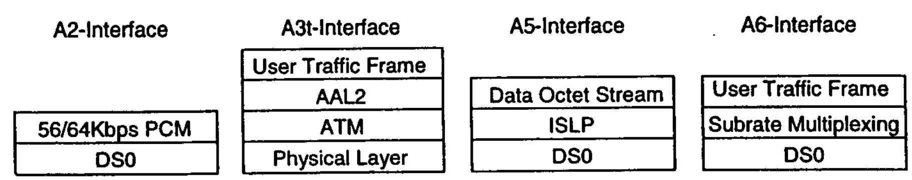


Figure 12.26 Transport protocol options for user traffic connection.

12.7.3.3 Data Service Protocols

The data service protocols for IS-95 are specified in TIA/EIA/IS-707 "Data Service Options for Wideband Spread Spectrum Systems." The IS-707.4 specifies asynchronous

fax and data services, and IS-707.5 specifies packet data services. The L-interface is specified in TIA/EIA/IS-658.

Figure 12.27 shows an overview of data service protocols based on IS-658. The physical layer in the network side is based on B-channels or H-channels with unrestricted digital information as defined in ANSI T1.607-1990, or an ANSI/IEEE 802.3 LAN. FR SVC is frame relay switched virtual circuit, which conforms to ANSI T1.618. The adaptation layer accepts data octets from the link layer and the relay function, assembles them into blocks of data octets for transmission in the information field of FR SVC frames, or vice versa, accepts blocks of data octets, and disassembles them. Radio link protocol (RLP) reduces the error rate by applying retransmission protocol for erroneous packets. RLP is specified in 707.2.

The link layer is based on the point-to-point protocol (PPP), defined in RFC 1661. PPP provides a multiplexed method to carry higher layer protocols over serial links. Above PPP, asynchronous data and fax data applications use link control protocol (LCP) and internet protocol control protocols (IPCP). LCP provides a mechanism for the mobile station and the IWF to negotiate various options provided by PPP. IPCP allows the mobile station to request a temporary IP address from the IWF.

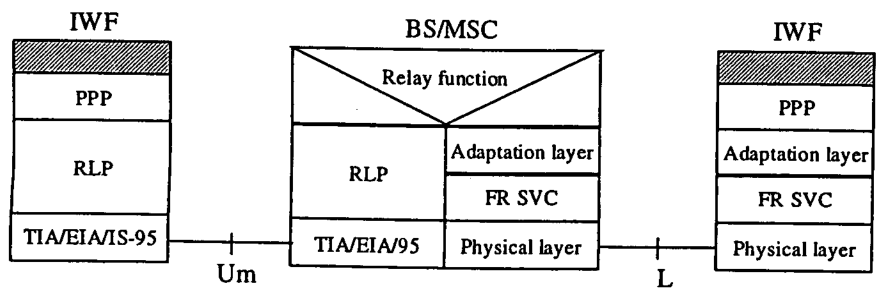


Figure 12.27 Mobile data protocol overview. (*Source*: TIA/EIA/IS-658, reproduced under written permission of the copyright holder (Telecommunications Industry Association).)

IS-707.5 specifies two protocol options for packet services: relay layer R_m interface protocol option and network layer R_m interface protocol option. The network layer R_m interface protocol option supports Terminal Equipment (TE) applications where the mobile termination is responsible for aspects of packet mobility management and network address management. The relay layer R_m interface protocol option supports TE applications where the TE is responsible for aspects of packet mobility management and network address management.

12.7.4 Evolution of IS-41/IS-95 Towards Third Generation

The standardization work of IS-41/IS-95-based networks towards third generation has recently started. In this section a few aspects relevant to the evolution are discussed.

12.7.4.1 *Architecture Evolution*

The ATM transmission infrastructure will provide efficient support for transmission of bursty wideband services. IS-634 is already using ATM in BS-to-BS communication. However, this needs to be extended into the A-interface towards MSC and possibly to the L-interface.

12.7.4.2 *Protocol Evolution*

Since the IS-41 or IS-95 were not specified using some formal specification language, the partitioning between protocols is sometimes difficult to identify. Interactions between different protocols make their independent evolution difficult. Thus, a clearer separation between different protocol layers is required before introducing third generation enhancements.

The BSSAP protocol in IS-634 needs to be evolved in a similar manner as in GSM. In addition, IS-41 call processing and mobility management protocols can be reused with improvements. If global roaming is desired, modifications are necessary for IS-41. For example, interworking with GSM MAP should be developed. One improvement for IS-41 could be the specification of soft handover between different MSCs. To simultaneously support multiple traffic channels changes are required in IS-41.

Most likely the PPP layer will be common to all data services, as in the current IS-707 standard. However, below the PPP protocol, a MAC layer will be specified to. support packet and possibly circuit switched data services. Furthermore, the RLP protocol needs revisions to accommodate new frame sizes.

REFERENCES

[1] ITU-T Recommendation I.130, "Method for the Characterization of Telecommunication Services Supported by an ISDN Network Capabilities of an ISDN."

[2] ITU-T Recommendation X.200, "Information Technology – Open Systems Interconnection – Basic Reference Model: The Basic Model."

[3] ITU-T Recommendation I.320,"ISDN Protocol Reference Model."

[4] Mouly, M., and M.-B. Pautet, *The GSM System for Mobile Communications*, published by the authors, 1992.

[5] Haferbeck, R., L. Laitinen, and S. Huusko, "BSS Aspects on the Radio Dependent Functionality," Issue 1, *CEC Deliverable, Future Radio Wideband Mobile Access System – FRAMES*, November 30, 1997.

[6] ETSI, "Global Multimedia Mobility (GMM), A Standardization Framework," Part A and B, 1996.

[7] E. Berruto, G. Colombo, P. Monogioudis, A. Napolitano, and K. Sabatakakis, "Architectural Aspects for the Evolution of Mobile Communications Towards UMTS," *IEEE Journal on Selected Areas in Communications*, Vol. 15, No. 8, October 1997, pp. 1477–1487.

[8] ITU-T Recommendation Q.FIN, "Framework of IMT-2000 Networks."

[9] ITU-T Recommendation I.312/Q.1201, "Principles of Intelligent Network Architecture."

[10] Modarressi, A. R., and R. A. Skoog, "Signaling System No. 7: A Tutorial," *IEEE Communications Magazine*, July 1990, pp. 19–35.

[11] Lin, Y.-B., and S. K. DeVries, "PCS Network Signaling Using SS7," *IEEE Personal Communications*, Vol. 2, No. 3, June 1995, pp. 44–55.

[12] Perkins, C. E., *Mobile IP: Design Principles and Practises*, Reading MA: Addison-Wesley Publishing Company, 1998

[13] Perkins, C. E., "Mobile IP", *IEEE Communications Magazine*, May 1997, pp. 84–99.

[14] ETSI, UMTS 23.20, "Universal Mobile Telecommunications System (UMTS), Evolution of the GSM Platform towards UMTS."

[15] White, P. P., "RSVP and Integrated Services in the Internet: A Tutorial," *IEEE Communications Magazine*, Vol. 35, No.5, May 1997, pp. 100–106.

[16] ETSI, UMTS 23.70, "Universal Mobile Telecommunications System (UMTS); Service Aspects; Virtual Home Environment."

[17] GSM 03.60, "General Packet Radio Service (GPRS); Service Description; Stage 2 ETSI," GSM Technical Specification 03.60, Version 5.0.0, June 1997.[2]

[18] Cai, J., and D. J. Goodman, "General Packet Radio Service in GSM," *IEEE Communications Magazine*, October 1997, pp. 122–131.

[19] Hämäläinen, J., *Design of GSM High Speed Data Services*, Ph.D. Thesis, Tampere University of Technology, Finland, 1996.

[20] Jung, P., Z. Zvonar, and K. Kammerlander (eds.), *GSM: Evolution Towards 3rd Generation*, Boston: Kluwer Academic Publishers, 1998.

[21] ETSI, UMTS 23.30, "Universal Mobile Telecommunications System (UMTS) Iu Principles."

[22] Verkama, M., "CAMEL and Optimal Routing, Evolution of GSM Mobility Management towards UMTS," in *GSM: Evolution towards 3rd generation*, P. Jung, Z. Zonar, and K. Kammerlander (eds.), Boston: Kluwer Academic Publishers, 1998.

[23] Gallagher, M. D., and R. A Snyder, *Mobile Telecommunications Networking with IS41*, New York: McGraw-Hill, 1997.

[24] TIA/EIA/SP-3693, "Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular Systems," to be published as ANSI/TIA/EIA-95B.[3]

[2] The present technical specifications are the property of ETSI. No reproduction shall be made without ETSI authorization. The original version of these ETSI technical specifications can be obtained from the Publication Office of ETSI.

[3] To purchase complete version of any TIA document, call Global Engineering Documents at 1800-854-7179 or send a facsimile to 303-397-2740.

Chapter 13

SYSTEM COMPARISON

13.1 INTRODUCTION

This chapter presents comparisons between wideband CDMA and other third generation air interface candidates. The comparison criteria used can be divided into objective and subjective criteria. A comparison based on a single objective criterion is free of interpretations if the results have been obtained using the same assumptions for the different considered technologies. Examples of objective criteria are capacity and coverage. An evaluation based on subjective criteria is not free of interpretations. Examples of subjective criteria are the impact of the air interface in the fixed network and flexibility of the air interface. Very often criteria are also related to each other, for example, more sophisticated receiver techniques, such as interference cancellation, result in better performance but also in larger complexity. The overall judgment depends on the weight given to each criterion. Furthermore, a comparison is always a snapshot of the available technologies at a specific moment, and new information might change the judgment. It is also important to understand which technologies form an essential part of an air interface and which technologies are air interface independent. For example, speech coding is not specific to the air interface.

Due to the complex nature of the third generation systems, we do not try to present a comprehensive comparison of the proposed technologies. Rather, we explain and interpret the results of earlier comparisons. Based on this, we assess the feasibility of wideband CDMA for third generation air interface.

The chapter begins with a discussion of the comparisons performed for the second generation systems. Second, the SIG5 comparison results for ATDMA and CODIT air interfaces are presented. Next, a performance comparison of wideband TDMA and wideband CDMA for 2-Mbps transmission is described. The performance of wideband CDMA and GSM air interfaces is discussed. The FRAMES project compared different multiple access schemes, the results of which are presented here. A

short review of ARIB air interface comparison follows. Furthermore, we discuss the ETSI UMTS radio access (UTRA) scheme evaluation. Based on the review of the different air interface comparisons, we conclude with a discussion on the feasibility of wideband CDMA as a third generation air interface.

13.2 SECOND GENERATION AIR INTERFACE COMPARISONS

So far, the probably most well-known comparison of air interfaces, or rather the most well-known debate, has been between GSM and IS-95. Without going into the details, we can note that the published results from these comparisons are very controversial [1–5]. The main deficiency of these comparisons is that no common assumptions were specified, and thus, the obtained results were not directly comparable. Furthermore, since source codecs, not part of the air interface as such, were included, these comparisons do not actually reflect the performance differences between TDMA- and CDMA-based air interfaces. A more realistic comparison can be found in [5].

For the US PCS system selection, an extensive list of criteria was developed. The advantages and disadvantages of different air interfaces were compared based on these criteria. Originally, 16 proposals were submitted [6]. The merging of the system porposals was based on a voluntary consolidation, which resulted in reducing the number of air interfaces to seven. These are GSM1900, IS-95, IS-136, Digital European Cordless Telephone (DECT), Personal Access Communication System (PACS), W-CDMA, and a composite CDMA/TDMA.

13.3 RACE II - SIG5 COMPARISON

SIG5, a special interest group in the partly-EU funded RACE II program in Europe, evaluated the ATDMA and CODIT air interfaces [7]. CODIT is a wideband CDMA system described in Chapter 6. In the comparison, a 5-MHz system bandwidth was used. ATDMA is an advanced TDMA system with a bandwidth of 270.92307 kHz and 1107.692308 kHz in macro- and micro/picocell environments, respectively [8].

In order to achieve compatible results, SIG 5 specified a set of common assumptions [9]. Services used in the comparison were 12-Kbps speech service and 64-Kbps data services. However, since ATDMA and CODIT did not agree on these common assumptions in the beginning of the projects, it was not possible to unify all assumptions and models in SIG5. The specification of QoS for CODIT and ATDMA, as well as the system simulation tools, were different. Signaling overheads were calculated in different ways. In addition, the speech codecs of the systems were different, and thus, no real comparison could be performed for speech services.

Since the assumptions to generate spectrum efficiency results were not exactly the same, it was difficult to draw any definitive conclusions from the SIG5 performance comparison. However, it was noted that CODIT performed better for data services in a macrocell environment since it had stronger channel coding [7]. Furthermore, CODIT performed better for mixed services (speech and data) than ATDMA, mainly due to better interference averaging. In micro- and picocells, the system deployment had a large impact on the results. ATDMA performed better, but CODIT would have gained more from sectorized antennas. In a microcell environment, the isolation between cells is high and the intracell interference limits the CDMA performance. On the other hand, TDMA is orthogonal within a cell and, since the intercell interference is low, it achieves good spectrum efficiency. The gain from sectorized antennas in the CDMA case is translated into direct capacity gains due to the reduction of intracell interference, while in TDMA the reuse factor cannot be reduced in direct proportion to the number of sectors. The same conclusions apply for a picocell environment.

Later, an ATDMA and CODIT comparison was continued and reported in [10]. For this study, more detailed common assumptions were agreed upon. The results were similar to the SIG5 comparison (i.e., CODIT performed better for data services in macrocell). In addition, CODIT performed also better for 64-Kbps data service in a microcell environment. This difference compared to the SIG5 results was most likely due to slightly different deployment scenarios and assumptions used in the two studies.

In addition to the quantitative evaluation of performance, SIG5 evaluated the CODIT and ATDMA concept qualitatively using subjective criteria. A summary of this comparison is presented in Table 13.1. As can be seen, drawing any hard conclusions from this kind of subjective comparison is difficult. However, service flexibility seems to be the main advantage of CODIT.

13.4 TDMA VERSUS CDMA FOR 2 MBPS SERVICE

The performance of four wideband TDMA and one wideband CDMA air interfaces for 2-Mbps transmission has been evaluated in [11]. The details of the wideband TDMA schemes have been presented in [12] and the wideband CDMA schemes in [13]. Two of the TDMA schemes (2-MHz bandwidth with B-O-QAM and Q-O-QAM modulation) were the basis for the FRAMES FMA1 scheme, which was later adopted as the IS-136 indoor carrier (with a bandwidth of 1.6 MHz) [14]. The wideband CDMA scheme was later submitted to the FRAMES project, where it was adopted with some modifications as the FMA2 wideband CDMA scheme [14]. The FMA2 uplink specification was then adopted for the ETSI WCDMA uplink scheme. In this study, the wideband CDMA system had bandwidths of 5, 10, and 20 MHz.

The link level performance was simulated in macro-, micro-, and picocell environments. The cell capacity was simulated in micro- and macrocell environments. In addition, link budget (range) was analyzed for macro-, micro-, and picocells.

In general, the capacity and range of the TDMA and CDMA schemes were comparable, with just some minor differences. In the uplink microcell channel with 3 km/h mobile speed, the wideband CDMA performance with 5- and 10-MHz bandwidths was somewhat better than TDMA due to fast power control. In this channel, frequency diversity alone did not give full benefit. The downlink capacity of TDMA with reuse factor of one was slightly better than the CDMA capacity. The reuse factor of one was possible due to interference cancellation [15]. Without it, the reuse factor would have been larger and thus the capacity comparable to the wideband CDMA capacity.

Table 13.1
Summary of the Qualitative Comparison Between CODIT and ATDMA Concepts

	CODIT	ATDMA
Linear transmitter and receiver requirements	Linear modulation QPSK and OQPSK used	Linear modulation B-O-QAM and Q-O-QAM used
Handportable requirements – need for duplexer	Required	Required for higher bit rates
Power control issues	Open and closed loop Dynamic range in UL 80 dB for open loop and 12 dB for closed loop Command rate 2 Kbps (0.5 ms) Step size 1 dB	Open loop, optional fast loop in uplink Dynamic range 30 dB Command rate 80 – 320 ms or no power control Step size 1 dB
Dynamic channel allocation	N/A	Used to achieve higher capacity
Frequency hopping issues	N/A	Used, gain depends on correlation bandwidth and available bandwidth
Sharing frequency band capability	CDMA less sensitive to narrowband interference; Sharing very difficult due to near-far effect, might be possible if same cell sites or interference cancellation used	Sharing might be possible With FDMA or CDMA if CDMA power density low enough
Synchronization requirements for base station	Asynchronous; DL synchronization required during handover on call-by call basis	No synchronization required; Might be beneficial to speed up scanning and handover especially in micro- and picocells
Frequency and code planning	No frequency planning except for fitting CODIT frequency channels (1, 5, and 20 MHz) into the allocated bandwidth (e.g., 30 MHz) Pilot code channels need to be planned	Either fixed frequency plan or DCA; With DCA common control channels need to be planned.
Service flexibility	Universal transport layer Variable rate services frame-by-frame basis Simultaneous speech and data Connectionless data	Data rates for low delay and long delay constrained services: 9.6, 19.2, 64 Kbps, up to 2 Mbps Unconstrained delay data
Handover performance	Soft handover Interfrequency handover, using compressed mode results to 1.5 dB degradation in C/I ratio	Backward handover Forward handover PRMA++ allows reservation of resources in a new cell before releasing the old traffic channel Macro diversity specified, but not as beneficial as in CDMA
Distributed and adaptive antennas	Not studied	Not studied
EMC issues	CDMA has continuous transmission in reverse link, power control modifies very rapidly the total transmission power	Pulsed transmission

13.5 COMPARISON OF WIDEBAND CDMA AND GSM

The impact of different radio channel and service conditions on wideband CDMA and GSM air interfaces has been studied in [16]. The wideband CDMA air interface is the same as in [16], which was later used as the basis for the FMA2 scheme in FRAMES [17]. Evaluation of spectrum efficiency was carried out for speech, 12-Kbps data service, and 96-Kbps data service. Furthermore, a mixed service case with 80% speech and 20% 12-Kbps data users was evaluated. The required BER is 10^{-3} for all services. The evaluated services are summarized in Table 13.2.

Table 13.2
Evaluated Services

	Data rate	BER	Maximum delay (ms)
Speech (GSM)	6.5, DTX = 0.5	10^{-3}	60
WB-CDMA	8, DTX = 0.4	10^{-3}	30
Low rate data	12	10^{-3}	100
High rate data	96	10^{-3}	100

For both air interfaces, the spectrum efficiency of the downlink was less than in the uplink. To remove the imbalance between up and downlink, introduction of antenna diversity at the terminal should be considered. For speech and data services, the downlink of the GSM air interface provided equal performance compared to wideband CDMA. However, the uplink performance of wideband CDMA was better that the performance of GSM. This was due to multiuser detection. Furthermore, the reduction of E_b/N_0 due to antenna diversity used in the uplink turns into higher capacity gain for CDMA. In general, all radio access technologies can provide equal spectrum efficiency, as was shown by this study. However, by analyzing the performance of different services, critical design factors can be identified and thus further improvements developed. For GSM/DCS, the following learned lessons are summarized below:

- Interference cancellation can substantially improve the performance for both up and downlink.
- Burst-by-burst hopping improves performance of high data rate services.
- DTX gain does not always turn into capacity improvement due to nonlinear increase in interference as the reuse factor is reduced.
- The GSM half-rate codec introduces unnecessary delay and better interleaving scheme would improve performance.

The following lessons for wideband CDMA were learned:

- All energy from the channel needs to be collected to obtain good performance.
- Power control is essential to provide good performance.
- Uplink power control signaling reduces the spectrum efficiency of downlink especially for low bit rate speech service.

- In a channel with a large number of equal gain taps, soft handover does not provide additional diversity due to limited number of RAKE fingers in the mobile station.

13.6 FRAMES MULTIPLE ACCESS COMPARISON

The FRAMES project investigated hybrid multiple access technologies in order to select the best combination as a basis for further detailed development of UMTS radio access system. The multiple access evaluation results have been reported in [17]-[19]. The FRAMES evaluation campaign consisted of two stages. At the first stage, several candidate schemes were compared and schemes with similar characteristics were combined. Based on the results of the first stage, the multiple access schemes were divided into two groups:

- Multicarrier TDMA (multiples of 200 kHz), single wideband TDMA (WB-TDMA, bandwidth 1-2 MHz), and hybrid CDMA/TDMA (bandwidth 1.6 MHz);
- Asynchronous CDMA (WB-CDMA, bandwidth 6 MHz), OFDM/CDMA, and synchronous CDMA.

At the second stage, these schemes were evaluated against criteria derived from the UMTS requirements. Based on this evaluation, a harmonized multiple access platform was defined. The FRAMES multiple access (FMA) platform consists of two modes, FMA1 with and without spreading (bandwidth 1.6 MHz) and a wideband CDMA mode (bandwidth 6.4 MHz). In the following, we summarize the FRAMES evaluation of the FMA platform based on [17].

Provision of Various Data Rates in Different Environments and Bearer Service Flexibility. Both FMA1 and FMA2 can support the UMTS bit rates from low bit rates up to 2 Mbps. FMA1 supports variable bit rates with low granularity by DTX or resource re-assignment together with adaptive coding. The slotted structure of FMA1 suits well bursty packet type services. In FMA2, variable bit rates are supported by adaptive coding and power assignments. Due to this power sharing and long initial synchronization, FMA2 is more suitable for moderately varying circuit switched services. For both modes, different operating points for the services are obtained by different combinations of coding and link adaptation. In FMA1, mixed bearer services can be provided to one user by packing them into different slots/codes; while in FMA2, mixed bearer services for one user can be multiplexed and have different operating points. For both modes bearers/services can be added and dropped during a call.

Spectrum Efficiency (capacity). The spectrum efficiency for FMA1 is presented in Table 13.3 and for FMA2 in Table 13.4. The prevailing conditions for the spectrum efficiency simulations are summarized below:

- Hexagonal cell layout;
- Uniformly distributed mobile stations;
- 5% outage probability;
- Pathloss law with a decay factor of 3.6;
- 10 dB shadowing parameter;
- Power control used (except for FMA1 mixed services);
- Errors due to power control and handover taken into account;
- FMA1 with spreading evaluated (orthogonal codes);
- For FMA1 frequency reuse is optimized for each case;
- FMA2 had frequency reuse factor of 1.

Table 13.3
Spectrum Efficiency of FMA1

Service	Spectrum Efficiency (Kbps/MHz/cell), load/reuse factor	
	Downlink	Uplink
12 Kbps, BER 10^{-3}, 40 ms	124 , 77%/3	250, 52%/1
144 Kbps, BER 10^{-3}, 40 ms	126, 70%/4	240, 33%/7
Mixed services 90% / 10% 12 Kbps / 144 Kbps	57 (no power control)	95 (no power control)
2 Mbps, BER 10^{-3}, 100 ms	Downlink is limiting direction; Spectrum efficiency 100 - 150 Kbps/MHz/s.	

Table 13.4
Spectrum Efficiency of FMA2

Service	Spectrum Efficiency (Kbps/MHz/cell)	
	Downlink	Uplink (with multiuser detection)
Speech/low rate data 12 Kbps, BER 10^{-3}, 40 ms	108	192
Medium data 144 Kbps, BER 10^{-3}, 100 ms	108	389
Mixed services (12 Kbps/144 Kbps)	115	322
2 Mbps, BER 10^{-3}, 100 ms	Downlink is limiting direction Spectrum efficiency 100 - 150 Kbps/MHz/s.	

Spectrum efficiency figures in the downlink are very close to each other. Poor performance of mixed service in FMA1 is due to missing power control, which was not used due to modeling difficulties. For both modes, downlink is the limiting direction since uplink receiver antenna diversity was used. The better performance of FMA2 in the uplink is due to multiuser detection. Multiuser detection efficiency (i.e., the amount

of own-cell interference that can be removed) varies in different environments and here it was assumed to be 60%.

Coverage. Coverage evaluation was carried out for the same services and conditions as for spectrum efficiency. The access scheme -dependent parameters determining range are:

- Link level performance E_b/N_0;
- The amount of overhead transmission.

E_b/N_0 figures taking into account the overhead transmission due to power control and training bits are presented in Table 13.5. The differences are very small, and the uplink is the limiting direction due to lower transmission power. For the low bit rate services, FMA1 has a slight advantage; while for the 144-Kbps service, FMA2 has an advantage.

Table 13.5
E_b/N_0 results

	FMA1	
E_b/N_0	Downlink	Uplink
Speech/low rate data 12 Kbps, 10^{-3}, 40 ms	7.3 dB	3.4 dB
Medium rate data 144 Kbps, 10^{-3}, 40 ms	5.9 dB	4.4 dB
	FMA2	
E_b/N_0	Downlink	Uplink
Speech/low rate data 12 Kbps, 10^{-3}, 40 ms	6.9 dB	7.2 dB
Medium data 144 Kbps, 10^{-3}, 100 ms	6.5 dB	3.1 dB

Support for Adaptive Antennas. Both modes support adaptive antenna techniques. In FMA1, capacity gains are realized through smaller cluster sizes and in FMA2 by allocating more codes. SDMA can be used for C/I and capacity improvements. In FMA2, a user dedicated reference signal instead of a common pilot signal is needed in the downlink.

Hierarchical Cell Structures (HCS). Hierarchical cells are supported by interfrequency handover, which in FMA1 is an inherent part of the system design and in FMA2 is supported through discontinuous uplink transmission and dual receivers in the downlink. In both modes, the mobile terminal assists with measurements for handover (MAHO). Handover between FMA-based UMTS and second generation (GSM) is supported through dual-mode terminals together with the capability of measuring GSM BCCH frequencies from FMA, owing to the choice made above for the clock rate and frame structure.

Duplex Method, Support for Public and Private Environments. FMA1 is better suited for operation in private environments than FMA2, since it requires less coordination. FMA1 supports asymmetric data services in TDD mode since the number of slots between uplink and downlink can be varied. FMA2 could, in principle, operate in TDD as well, but the basic transmission scheme should be modified in that case, knowing also that flexible allocation of resources between up and downlink is difficult.

Terminal Impacts (Power Consumption and Complexity, GSM/UMTS Dual-Mode Terminals). In the evaluation of terminal impacts, it can be noted that cost, size, and power consumption for baseband complexity decrease drastically with time, enabling more complex baseband algorithms like multiuser detection and interference cancellation. RF power consumption is mainly determined by the power level and linearity requirements of the output amplifier. In macrocells with low bit rates, RF power consumption surpasses baseband power consumption. In FMA1, wide bandwidth and high data rates give stringent A/D converter requirements. In addition, higher order modulation and multicode transmission increase the power amplifier linearity requirements. FMA2 power amplifier requirements are less stringent due to variable spreading gain in the uplink. In FMA1, a dual receiver is needed for inter-frequency handovers at higher data rates, while in FMA2 a dual receiver is always needed to support inter-frequency handover. A dual-mode UMTS and GSM/DCS terminal needs additional GSM/DCS duplexer and RF filters, but it can reuse the same reference oscillator for both modes with a careful parameter design.

BSS Impacts (Evolution from Existing Systems, BSS Complexity and Cost). Regarding BSS impacts, FMA1 operates with hard handover. In FMA2, soft handover is required, necessitating additional costs due to 1.5 times more transmission capacity plus diversity combining. In FMA1, integration of new cells can be handled by slow dynamic channel allocation (DCA) to simplify network planning, and in FMA2 by automatic power planning.

13.7 ARIB COMPARISON

ARIB has reported a performance comparison between TDMA (MTDMA system), OFDM (BDMA system), and CDMA technologies [21]. The comparison was based on simulations and experimental systems. The conclusions were as follows [21]:

- Experiment data suggest that the CDMA system has a good feasibility although some details require further study.
- The major technical contention of MTDMA system is high bit rate transmission. MTDMA could be an alternative to the CDMA system in office and pedestrian environments.
- BDMA offers potentiality for capacity expansion. However, it is difficult to judge at this stage, due to the lack of data based on theoretical studies and experimental verification.

Later, it was decided not to adopt BDMA and MTDMA for further standardization. The advantages of wideband CDMA for IMT-2000 were stated to be flexible high data rate services along with high quality, improved multipath resolution and increased interference averaging. Time-to-market aspects could also be counted as advantages for the CDMA solution. No detailed technical results are publicly available from the ARIB comparison.

13.8 ETSI AIR INTERFACE COMPARISON

During 1997, ETSI SMG2 performed an air interface concept evaluation in order to select the UMTS terrestrial radio access (UTRA) scheme. Four concepts were compared in this process, namely, wideband CDMA (concept α), OFDM (concept β), wideband TDMA (concept γ), and TD-CDMA (concept δ). The spectrum efficiency results are summarized in Table 13.6, and the qualitative comparison results in Table 13.7. LCD means low constrained delay data and UDD unconstrained delay data (i.e., packet data). The evaluation criteria, common assumptions, and models are described in [22]. The evaluation results have been presented in [23].

The level of evaluation details for different concepts is highly variable. However, it can be concluded that all schemes seem to satisfy the criteria. Wideband CDMA seems to be the most flexible for variable bit rate and mixed service applications. However, wideband CDMA requires the largest minimum spectrum allocation. The spectrum efficiency of wideband CDMA was best for speech and 384-Kbps services in the downlink. In an indoor office, wideband TDMA had comparable spectrum efficiency to wideband CDMA. Furthermore, downlink antenna diversity improved the spectrum efficiency of all schemes considerably.

In general, all schemes evaluated in ETSI could have formed the basis for UMTS air interface. However, wideband CDMA was selected for FDD due to its technical merits, such as flexibility and high spectrum efficiency, as well as its international support and the large amount of research that had resulted in a detailed specification. This was judged to reduce the risk for commercial systems. TD-CDMA was selected for TDD, as a result of a compromise, since it had the largest support after wideband CDMA in the final voting. Furthermore, as discussed in Chapter 9, wideband CDMA TDD mode naturally contains time division principle, and thus, the political compromise was also viable from a technical point of view.

13.9 CONCLUSIONS

What is the best or the most suitable air interface for third generation wireless communication systems? From a purely technical point of view, there is no definitive answer. Each scheme has some strong and weak points. In addition, if there are differences, it is a very subjective matter to determine how significant they are.

In general, we can state that any of the proposed technologies, TDMA, CDMA, OFDM, or hybrid CDMA/TDMA could be used as an IMT-2000 air interface. The technical differences are not that large. Furthermore, we can state that many of the

Table 13.6
Spectrum Efficiency Results in Kbps/MHz/cell for UTRA Concepts

Environment	Service	WCDMA (concept α) Uplink/downlink	OFDMA (concept β) Uplink/downlink	WB-TDMA (concept γ) Dl results only since limiting direction	WB-TD-CDMA (concept δ) Dl results only since limiting direction
Vehicular	Speech	98 / 78	33/31	55	72 (FH) 68 (no FH)
	LCD384	138/85 204/123 (30 dBm MS, 8 Mcps) 138/211 175/211 (30 dBm MS) 204/250 (30 dBm MS, 8 Mcps)	152/208 (with optimized transmitter power)	113	129/176 176 (DL diversity)
Outdoor-to-indoor pedestrian	UDD384	470 / 565	440/465	811	812 (DL ant div, reuse 1) 387 (DL ant div, reuse 3)
	Speech	127 / 163 189 / - (C/I based HO)	30.75/32.25	190	75 (FH) 73 (no FH)
Indoor office	service mix - speech - UDD384	315 / 207 315 / 460 (DL ant div)	TBD	60	110 (FH) 104 (no FH)
	UDD2048	300 / 230 300 / 500 (DL ant div)	240/240	332 743 (wall attenuation 5 dB)	170 (reuse 1) 195 (reuse 3) 405 (DL ant div, reuse 1) 132 (DL ant div, reuse 1)

Source: [23].

noticed weaknesses of the previous schemes can be overcome by further development. For example, in the FRAMES comparison, CDMA was noted to be better suited for moderately varying bit rates rather than for packet data. Since then, however, a new packet access mode for WCDMA in ETSI and ARIB has been developed (see Section 6.3.6) and this drawback has been circumvented. In FRAMES, no antenna diversity was used in the downlink, and thus, it was concluded that the downlink limits capacity. However, different solutions such as antenna diversity in the mobile station, base station transmit diversity, and interference cancellation in the mobile station are being discussed as a solution for this problem.

Table 13.7
Summary of the Comparison Results for the UTRA Concepts

	WCDMA (concept α)	OFDMA (concept β)	WB-TDMA (concept γ)	WB-TD-CDMA (concept δ)
Bearer capabilities • Rural outdoor: at least 144 Kbps (goal to achieve 384 Kbps), maximum speed: 500 km/h • Suburban outdoor: at least 384 Kbps (goal to achieve 512 Kbps), maximum speed: 120 km/h • Indoor/low range outdoor: at least 2Mbps, maximum speed: 10 km/h • UTRA should allow evolution to higher bit rates.	Supported with 4.096 Mcps chip rate With 8.196 and 16.392 Mcps bit rates up to 4 and 8 Mbps supported	Supported In indoor, over 2 Mbps supported	Supported Up to 4 Mbps for short range	Supported Higher bit rates can be supported using higher order modulation or higher RF bandwidth
Flexibility • Negotiation of bearer service attributes • Parallel bearer services, real-time/non real-time communication modes • Adaptation of bearer service bit rate • Circuit and packet oriented bearers • Supports scheduling of bearers according to priority • Adaptation of link to quality, traffic, and network load, and radio conditions • Wide range of bit rates should be supported with sufficient granularity • Variable bit rate real time capabilities should be provided • Bearer services appropriate for speech shall be provided	Supported Bit rates from 100 bps up to 2 Mbps with a granularity of 100 bps, change of bit rate on 10-ms basis TDD mode can be used for asymmetric services	Supported Granularity of 13 Kbps, finer granularity by varying channel coding TDD mode can be used for asymmetric services	Supported Two different size time-slots, granularity by varying channel coding TDD mode can be used for asymmetric services	All supported Variable bit rates provided by pooling of time slots and CDMA codes, granularity by varying channel coding TDD mode can be used for asymmetric services
Handover • Provide seamless (to user) handover between cells of one operator • The UTRA should not prevent seamless HO between different operators or access networks • Efficient handover between UMTS and second generation systems (e.g., GSM) should be possible	Supported Soft handover and inter-frequency handover with slotted mode for single receiver mobile stations Either by dual receiver or by slotted mode for GSM carrier measurements WCDMA and GSM have same multi-frame structure	Supported Hard handover Supported Same frame structure allows synchronization of UMTS and GSM cell sites for easier handover	Supported Hard handover Supported Same frame structure allows synchronization of UMTS and GSM cell sites for easier handover	Supported Hard handover Supported Same frame structure allows synchronization of UMTS and GSM cell sites for easier handover

Table 13.7 (continued)

	WCDMA (concept α)	OFDMA (concept β)	WB-TDMA (concept γ)	WB-TD-CDMA (concept δ)
Compatibility with services provided by present core transport networks • ATM bearer services, GSM services, IP based services, and ISDN services	Supported by all schemes			
Radio access planning • If radio resource planning is required automatic planning shall be supported	In general, all schemes need coverage planning and planning of adequate capacity for different services (see Chapter 11). The overall planning effort is expected to be the same for all schemes			
Public network operators • It shall be possible to guarantee predetermined levels of QoS to public UMTS network operators in the presence of other authorized UMTS users.	To guarantee predetermined QoS levels, UMTS public operators require dedicated frequency bands with appropriate guardbands. Therefore, this requirement was not supported			
Private and residential operators • The radio access scheme should be suitable for low-cost applications where range, mobility, and user speed may be limited • Multiple unsynchronized systems should be able to successfully co-exist in the same environment • It should be possible to install base stations without coordination • Frequency planning should not be needed	Frequency avoidance techniques (e.g., not make an access on a frequency that is too disturbed) Power control is used to minimize interference but is still able react on increased received interference Multi-user detection and interference cancellation techniques Spectrum sharing between TDD WCDMA and FDD WCDMA systems possible	DCA, frequency hopping for interference averaging	DCA, interference averaging by frequency and time hopping	DCA, TX power limitations for private system
Variable asymmetry of total band usage • Variable division of radio resource between uplink and downlink resources from a common pool (NB: this division could be in either frequency, time, or code domains)	This is primarily supported by the TDD mode in all schemes. TDD modes can have variable switching point to provide asymmetry (see Chapter 9). For FDD, it is possible to pair different uplink and downlink portions by using a variable duplex distance CDMA based scheme can have a slightly easier trade-off between coverage and bit rate between up and downlink			

Table 13.7 (continued)

	WCDMA (concept α)	OFDMA (concept β)	WB-TDMA (concept γ)	WB-TD-CDMA (concept δ)
Spectrum utilization • Allows multiple operators to use the band allocated to UMTS without coordination[1] • It should be possible to operate the UTRA in any suitable frequency band that becomes available such as first and second generation system's bands.	Sharing supported in limited scenarios Carrier spacing between UMTS operators 4.4 to 5 MHz Reframing requires 5.2 MHz (GSM on one side and UMTS on the other band-edge) or 5.6 MHz if GSM is on both sides	DCA	DCA	Minimum spectrum for reframing is 3×1.6 MHz + guardbands, hot spot reframing with 1.6 MHz + guardbands
Coverage, capacity • System should be flexible to support a variety of initial coverage/capacity configurations and facilitate coverage/capacity evolution • Flexible use of various cell types and relations between cells (e.g., indoor cells, hierarchical cells) within a geographical area without undue waste of radio resources • Ability to support cost effective coverage in rural areas	Supported HCS supported within 15 MHz	Supported Flexible resource allocation	Supported DCA	Supported HCS supported with at least 3 layers, DCA can be used, also between layers
Mobile terminal viability • Handportable and PCMCIA card size UMTS terminals should be viable in terms of size, weight, operating time, range, effective radiated power, and cost	No major difference between technologies. This depends also on future developments			
Network complexity and cost • The development and equipment cost should be kept at a reasonable level, taking into account the cost of cell sites, the associated network connections, signaling load, and traffic overhead (e.g., due to handovers)	All supported			

[1] Spectrum sharing, without any coordination, in the same geographical area and still guarantee a level of quality of service to the users is impossible in any system. This requirement is for further study.

Table 13.7 (continued)

	WCDMA (concept α)	OFDMA (concept β)	WB-TDMA (concept γ)	WB-TD-CDMA (concept δ)
Mobile station types • It should be possible to provide a variety of Mobile Station types of varying complexity, cost, and capabilities in order to satisfy the needs of different types of users	No real differences			
Alignment with IMT-2000 • UTRA shall meet at least the technical requirements for submission as a candidate technology for IMT 2000 (FPLMTS)	All meet.			
Minimum bandwidth allocation • It should be possible to deploy and operate a network in a limited bandwidth	Minimum bandwidth 5 MHz including guardbands. For cosited UMTS operators 4.4 MHz carrier spacing can be used	Minimum bandwidth 5 MHz (no detailed guardband analysis) 0.8 – 1.6 MHz in single isolated cell, small network with 3.2 MHz	Minimum bandwidth 3×1.6 MHz.	Minimum bandwidth 3×1.6 MHz + guardbands
Electromagnetic compatibility • The peak and average power and envelope variations have to be such that the degree of interference caused to other equipment is not higher than in today's systems	Continuous transmission improves the peak-to-average power ratio and envelope variations compared with GSM and similar TDMA-based systems	Similar to GSM	Similar to GSM	Similar to GSM. Multicode and 16QAM modulation cause additional envelope variations
RF radiation effects • UMTS shall be operative at RF emission power levels which are in line with the recommendations related to electromagnetic radiation	In principle the average power levels are independent of RTT			
Security • The UMTS radio interface should be able to accommodate at least the same level of protection as the GSM radio interface does	From a ciphering point of view, all radio interface technologies offer the same level of protection as good as of GSM			

Table 13.7 (continued)

	WCDMA (concept α)	OFDMA (concept β)	WB-TDMA (concept γ)	WB-TD-CDMA (concept δ)
Coexistence with other systems • The UMTS terrestrial radio access should be capable to co-exist with other systems within the same or neighboring band, depending on systems and regulations	See spectrum utilization	No detailed guardband analysis done	No detailed guardband analysis done	No detailed guardband analysis done
Dualmode terminals • It should be possible to implement dual mode UMTS/GSM terminals cost effectively	To measure the GSM carriers in WCDMA mode, a slotted mode has been defined GSM RX RF and IF filters required	Harmonized clocks GSM RX RF and IF filters required	Harmonized clocks GSM RX RF and IF filters required	Harmonized clocks GSM RX RF and IF filters required

REFERENCES

[1] Qualcomm, "A Comparison of the Six C's of Wireless Communications: Coverage, Capacity, Cost, Clarity, Choice, Customer Satisfaction," *Technical Report*, Qualcomm, 80-12589, Revision x1, 1994.

[2] Mohr, W., and J. Farjh, "GSM and CDMA, an Objective Comparison," *GSM MoU/ECTEL MRSG, Latin American Seminar*, Buenos Aires, Argentina, December 4-5, 1995.

[3] Ericsson, "PCS 1900 The New Personal Communications System for North America," *Technical Report*, Ericsson, December 12, 1994.

[4] Ericsson, "IS-95 and PCS-1900 Capacity Comparison," *Technical Report*, Ericsson, July 6, 1994.

[5] Fagen, D., A. Aksu, and A. Giordano, "A Case Study of CDMA and PCS-1900 Using the GRANET Radio Planning Tool," *Proceedings of ICPWC*, Mumbai, India, December 1997, pp. 505–509.

[6] Cook, C. I., "Development of Air Interface Standards for PCS," *IEEE Personal Communications*, Vol. 1. No. 1, 1994, pp. 30–34.

[7] Pizarroso, M., and J. Jiménez (eds.), "Preliminary Evaluation of ATDMA and CODIT System Concepts," *SIG5 Deliverable MPLA/TDE/SIG5/DS/P/002/b1*, September 1995.

[8] Urie, A., M. Streeton, and C. Mourot, "An Advanced TDMA Mobile Access System for UMTS," *IEEE Personal Communications*, Vol. 2, No. 1, February 1995, pp. 38–47.

[9] Pizarroso, M., and J. Jiménez (eds.),"Common Basis for Evaluation of ATDMA and CODIT System Concepts," *SIG5 deliverable MPLA/TDE/SIG5/DS/P/002/b1*, September 1995.

[10] Barberis, S., P. Blanc, L. Bonzano, E. Gaiani, V. Palestini, and G. Romano, "Radio Transmission and Capacity Comparison between ATDMA and CODIT Systems for UMTS," *Proceedings of VTC'97*, Vol. 2, Phoenix, USA, May 1997, pp. 815–819.

[11] Pehkonen, K., H. Holma, I. Keskitalo, E. Nikula, and T. Westman "A Performance Analysis of TDMA and CDMA Based Air Interface Solutions for UMTS High Bit Rate Services," *Proceedings of PIMRC'97*, Helsinki, Finland, September 1997, pp. 22–26.

[12] Nikula, E., and E. Malkamäki, "High Bit Rate Services for UMTS Using Wideband TDMA Carriers," *Proceedings of ICUPC'96*, Vol. 2, Cambridge, Massachusetts, USA, September/October 1996, pp. 562–566.

[13] Westman, T., and H. Holma, "CDMA System for UMTS High Bit Rate Services," *Proceedings of VTC'97*, Phoenix, Arizona, USA., May 1997, pp. 824–829.

[14] Ojanperä, T., and R. Prasad, "Overview of Air Interface Multiple Access for IMT-2000/UMTS," *IEEE Communications Magazine*, 1998.

[15] Ranta, P., A. Lappeteläinen, and Z.-C. Honkasalo, "Interference cancellation by Joint Detection in Random Frequency Hopping TDMA Networks," *Proceedings of*

ICUPC'96, Vol. 1, Cambridge, Massachusetts, September/October 1996, pp. 428–432.

[16] Ojanperä, T., P. Ranta, S. Hämäläinen, and A. Lappeteläinen, "Analysis of CDMA and TDMA for 3rd Generation Mobile Radio Systems," *Proceedings of VTC'97*, Vol. 2, Phoenix, Arizona, USA., May 1997, pp. 840–844.

[17] Ojanperä, T., J. Sköld, J. Castro, L. Girard, and A. Klein, "Comparison of Multiple Access Schemes for UMTS," *Proceedings of VTC'97*, Vol. 2, Phoenix, Arizona, USA, May 1997, pp. 490–494.

[18] Ojanperä, T., K. Rikkinen, H. Häkkinen, K. Pehkonen, A. Hottinen, and J. Lilleberg, "Design of a 3rd Generation Multirate CDMA System with Multiuser Detection, MUD-CDMA," *Proceedings of ISSSTA'96*, Vol. 1, Mainz, Germany, 1996, pp. 334–338.

[19] Ojanperä, T., M. Gudmundson, P. Jung, J. Sköld, R. Pirhonen, G. Kramer, and A. Toskala, "FRAMES - Hybrid Multiple Access Technology," *Proceedings of ISSSTA96*, Vol. 1, Mainz, Germany, 1996, pp. 320–324.

[20] Ojanperä, T., P.-O. Anderson, J. Castro, L. Girard, A. Klein, and R. Prasad, "A Comparative Study of Hybrid Multiple Access Schemes for UMTS," *Proceedings of ACTS Mobile Summit Conference*, Vol. 1, Granada, Spain, 1996, pp. 124–130.

[21] Ishida, Y., "Recent Study on Candidate Radio Transmission Technology for IMT-2000," *First Annual CDMA European Congress*, London, UK, October 1997.

[22] ETSI UMTS 30.03, "Selection procedures for the choice of radio transmission technologies of the Universal Mobile Telecommunications System (UMTS)," *ETSI Technical Report*, v. 3.0.0, May 1997.

[23] ETSI UMTS 30.06, "UMTS Terrestrial Radio Access Concept evaluation," *ETSI Technical Report*, to be published by ETSI, 1998.

Chapter 14

STANDARDIZATION WORK AND FUTURE DIRECTIONS

14.1 INTRODUCTION

In Chapter 1, we divided the chronology of CDMA development into three eras: the CDMA pioneer era, the narrowband CDMA era, and the wideband CDMA era. The wideband CDMA era began with research activities in the early 1990s; it is now continuing with the adoption of wideband CDMA by several standards bodies. So far, ETSI, ARIB, and TR45.5 have defined wideband CDMA frameworks. This is only a starting point for detailed standardization. In order to move from the standardization phase into the implementation and commercial operations phases, detailed standards need to be ready. Furthermore, after a standard has been specified, a number of other elements need to be investigated and finalized in order to deploy commercial systems: spectrum needs to be acquired, applications developed, markets analyzed. Otherwise, a standard just gathers dust on a bookshelf among numerous other unused standards.

This chapter addresses the development of wideband CDMA proposals into commercial standards. We first describe the structure of standards bodies and industrial interest groups related to the third generation standardization in different regions. This establishes the basis for the discussion on third generation radio and network standardization. We address the role of regional standard bodies and the ITU in this process. Special attention is given to the ITU radio transmission technology (RTT) evaluation process. Regulatory developments such as spectrum licensing are described.

The third generation market, and the role of second generation systems in establishing it, are discussed. The roles of other technologies such as wireless LANs, wireless ATM, and satellites, which can either complement or compete with IMT-2000, are presented. We conclude this chapter by discussing enabling technologies and by describing a third generation service/application scenario.

14.2 STANDARDIZATION BODIES AND INDUSTRY GROUPS

This section introduces the different standardization bodies for the third generation systems, their responsibilities, and their organization. We start by describing the IMT-2000 standard organization in the ITU. Next, the standards bodies and the third generation industry interest groups in Europe, the United States, Japan, Korea, and China are presented. Furthermore, each subsection starts with a description of the general approach each region is taking to the development of mobile radio systems.

14.2.1 ITU

Figure 14.1 shows the organization of ITU for the IMT-2000 standardization. The telecommunications standardization in the ITU is divided into two sectors: Radiocommunications Sector (ITU-R) and Telecommunication Standardization Sector (ITU-T). The Intersector Co-ordination Group (ICG) is coordinating the IMT-2000 radio and network standards. The IMT-2000 radio aspects are being standardized in the Task Group (TG) 8/1 in ITU-R. TG8/1 is also responsible for the overall system architecture for IMT-2000 within ITU. In ITU-T, SG8 has been identified as the lead study group for coordination of standardization activities on IMT-2000 network standards. For more information on the ITU standardization see [1].

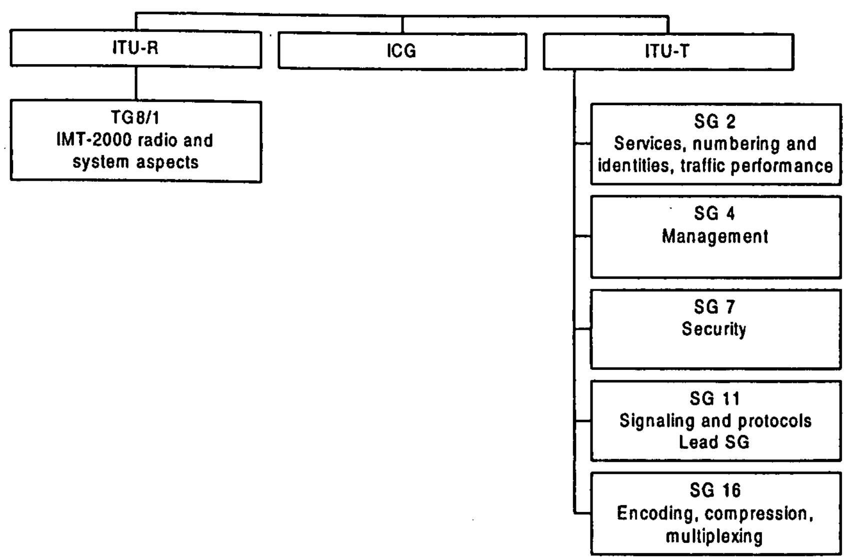


Figure 14.1 IMT-2000 standardization structure in ITU.

The role of the ITU in the standardization of third generation systems has been under discussion several times. The work in TG8/1 started in 1986. The telecommunications environment has changed dramatically since then. Therefore, it has been necessary to consider ITU's role with respect to regional standards bodies. The creation of the ITU Family of Systems concept has marked a clear shift in the ITU role (see Section 12.5.2). It has been recognized that second generation standards will evolve further and most likely form the basis for the third generation networks. Given the diverse need and the pace of development in different regions, a single global approach does not seem to be adequate. Furthermore, commercial interests due to the current investments into the second generation networks are further obstacles towards a single third generation standard. Based on these considerations, the role of the ITU seems to move towards setting global frameworks for requirements, spectrum allocations, interworking, and so on, rather than to the detailed drafting of standards.

14.2.2 Europe

In Europe, telecommunication standards are developed by ETSI. The general approach in Europe has been building consensus with emphasis on long-term solutions. The European Union is involved in the standards setting process by developing enabling regulatory policies. The global success of GSM is partly based on this type of standardization approach, and the Europeans are trying to carry this over to the third generation.

The ETSI structure related to UMTS developments is depicted in Figure 14.2. The most relevant technical committee (TC) from the UMTS point of view is TC SMG (Special Mobile Group). TC SMG is responsible for GSM and UMTS development. The actual standardization work is carried out in subtechnical committees. SMG2 has been responsible for the UMTS air interface evaluation and will standardize UMTS radio access network. SMG3 is responsible for GSM/UMTS core network standardization, and SMG12 is responsible of the overall UMTS architecture. SMG5, the original subtechnical committee responsible for UMTS, has been discontinued in 1997. TC NA (Network Aspects) has some network activities related to UMTS. TC SES (Satellite) is partly responsible for standardization of the UMTS satellite component. ETSI Project (EP) Digital Enhanced Cordless Telecommunication (DECT) is also considering some UMTS aspects. The work in TCs and STCs is based on the voluntary effort from ETSI members. In addition to STCs, a permanent nucleus, the Special Task Force, is supporting the TC SMG. For more information on the ETSI standardization see [2].

The UMTS Forum is a nonprofit association under the Swiss law. The UMTS Forum was established in 1996 on the recommendation of the UMTS Task Force [3]. It provides advice and recommendations to the European Commission, European Radiocommunications Office (ERC), ETSI, and national administrations. The UMTS Forum is divided into working groups (e.g., WG1 Regulatory, WG2 spectrum, and WG3 Market Aspects). For more information on the UMTS Forum see [4].

The GSM MoU Association is the GSM operators' organization for promoting and evolving the GSM cellular platform world-wide. Regional interest groups have been established to take requirements and special needs from the various regions into account: Asia Pacific Interest Group (APIG), Arab Interest Group (AIG), European Interest Group (EIG), North American Interest Group (NAIG), Central/Southern African Interest Group (CSAIG), East Central Asia Interest Group (ECAIG), and India Interest Group (IIG). In addition to GSM operators, government regulators/administrations can be members of the GSM MoU Association. Within GSM MoU, the Third Generation Interest Group (3GIG) was formed in 1994. The 3GIG is responsible for conveying GSM MoU vision and requirements into the standardization of UMTS. Furthermore, it addresses third generation market aspects and licensing issues. For more information on the GSM MoU see [5].

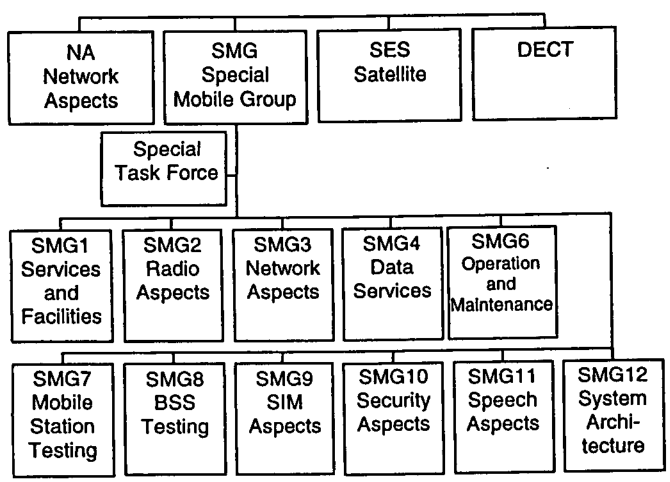


Figure 14.2 Partial view of the ETSI organization relevant to UMTS.

14.2.3 The United States

In the United States, the standards setting process is more driven by market inputs and industry interests than in Europe. On one hand, this leads to faster reaction to changes, but on the other, it leads to a larger number of standards. The selection between standards is then left to the market forces. The market-driven approach is reflected also in the standardization bodies. There are two main standardization bodies for mobile radio systems. The Telecom Industry Association (TIA) is a trade organization that provides its members with numerous services including standardization. The committee T1 on telecommunications was established in 1984 as a consequence of the breakup of

the Bell System [6]. Both the TIA and T1 can set American National Standards Institute (ANSI) accredited standards. Within TIA, the Wireless Communications Division is responsible for standardization of wireless technologies. The main committees from a third generation point of view are TR45 and TR46 (Public Mobile and Personal Communications Standards). Within T1, subcommittee T1P1 is responsible for management activities for personal communications systems (PCS). For more information on the TIA standardization activities see [7], and on T1P1 activities see [8].

14.2.3.1 TR45

Figure 14.3 shows the structure of TR45. In consists of six permanent subcommittees and six ad-hoc groups not shown in the figure. The responsibilities of the subcommittees are as follows:

- TR45.1 is responsible for the AMPS standardization.
- TR45.2 is responsible of the development of the IS-41 standard.
- TR45.3 is responsible for IS-136, the US TDMA system, and its evolution to UWC-136.
- TR45.4 is responsible for the IS-634 standard (i.e., the A-interface standard between BS and MSC). For the third generation, TR45.4 will most likely develop the interface between the access and core networks.
- TR45.5 is responsible for IS-95 technology and its evolution to cdma2000.
- TR45.6 work items include support for packet data and Internet access. TR45.6 coordinates its activities with TR45.2 to define network architecture for IMT-2000 that complies with packet data requirements.

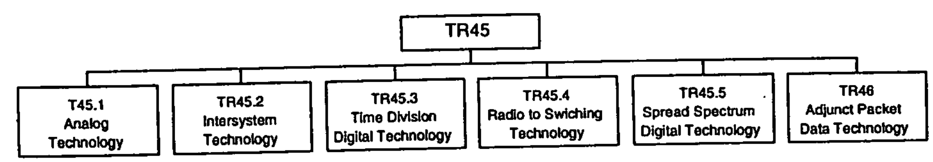


Figure 14.3 Structure of TIA TR45.

TR45.5 consists of four subcommittees as shown in Figure 14.4. TR45.5.1 is responsible for user needs and services such as multimedia services, speech, feature requirements, and data services. TR45.5.2 covers signaling and protocols including authentication/privacy, intersystem issues, call processing, and protocol tests. TR45.5.3 is responsible for the physical layer and has five working groups: RF parameters, supervision and malfunction detection, handoff and system timing, MS minimum performance, and BS minimum performance. TR45.5.4 is developing alternative

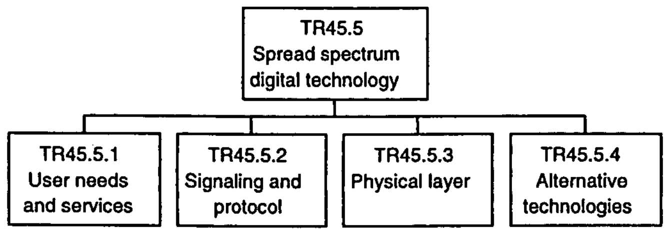


Figure 14.4 Structure of TIA TR45.5.

technologies and has been responsible for development of the framework for Wideband cdmaOne.

14.2.3.2 TR46 and T1P1

TR46 was created to address the TIA standards development for the US PCS. The main subcommittees within TR46 are listed below and represented in Figure 14.5.

- TR46.1 WIMS (Wireless ISDN and Multimedia Service);
- TR46.2 Cross technology;
- TR46.5 PCS1900;
- TR46.6 Composite CDMA/TDMA.

Standards committee T1P1 – Wireless/Mobile Services and Systems belonging to T1 organization – is closely related to TR46. They have an almost similar structure and jointly manage the above-mentioned TR46 standards. T1P1 has a cooperative arrangement with ETSI SMG regarding PCS1900/GSM technology.

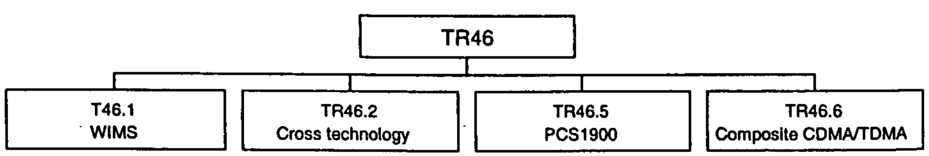


Figure 14.5 Structure of TIA TR46.

14.2.3.3 Industry Interest Groups

Each of the main PCS standards in the United States has its own industry interest group: the CDMA Development Group (CDG) for IS-95, Universal Wireless Communications Consortia (UWCC) for IS-136, GSM North America (GSM MoU regional interest group), and GSM Alliance. The last two are consortia for North American PCS1900 operators using GSM-technology.

CDG is an industry consortium of companies developing IS-95. The CDG is comprised of IS-95 service providers and manufacturers of subscriber and network infrastructure equipment. For more information on the CDG activities see [9]. The CDG organizational structure consists of the following entities [9]:

- Executive Board: a carrier-only board that oversees the CDG process. Participation on the board is limited to chief technical officers, chief operating officers, and chief executive officers.
- Steering Committee: a committee consisting of executives from member organizations. The objective of this committee is to review and approve the work of the supporting teams and address resource issues.
- CDG Working Groups (Teams): teams consisting of carrier and manufacturer business and technical resources. Advanced systems group is focusing on third generation technologies.

The UWCC is a limited liability corporation in Washington state, established to support an association of carriers and vendors developing, building, and deploying products and services based on IS-136 TDMA and IS-41 wireless intelligent network (WIN) standards. The UWCC structure consists of a board of governors and three forums: Global TFMA Forum (GTF), Global Win Forum, and Global Operators Forum. For more information on the UWCC activities see [10].

14.2.4 Japan

In Japan, the telecommunication operators and manufacturers have traditionally moved as one single entity. NTT has been in the leading role, dominating the wireless standardization scene, and the others have followed NTT in unison. The selection of IS-95 by DDI and IDO, two Japanese wireless operators, for replacing the analog system was an unusual break from the tradition of following NTT [11]. It has been very difficult for international companies to participate in the closed standards setting. One further obstacle has been the language. For third generation, Japan has opened standardization in order to create a global standard. Instead of Japanese, English has been selected as the working language for the documentation of the third generation standards.

In Japan, the IMT-2000 standardization is divided between two standardization organizations: Association of Radio Industries and Businesses (ARIB) and Telecommunication Technology Committee (TTC). ARIB is responsible for radio

standardization and TTC for network standardization. Contributions to ITU-R and ITU-T are done through the Telecommunication Technology Council of the Ministry of Posts and Telecommunications (MPT). MPT has established a Study Group on Next-Generation Mobile Communication, which gives recommendations to TTC and ARIB.

The organization of IMT-2000 Study Committee within ARIB is described in Figure 14.6. The Coordination Group (CG) in the Standards Subcommittee is responsible for international coordination with other standards bodies such as TTA, ETSI, and TIA. The harmonization discussions concerning the WCDMA and cdma2000 have been carried out in this group. SWG1 is responsible for system description; SWG2 is responsible for the specification of the air interface layer 1; and SWG3 is responsible for the air interface specification for layers 2 and 3. The requirements and objectives are established in the Application Group. For more information on the ARIB activities see [12].

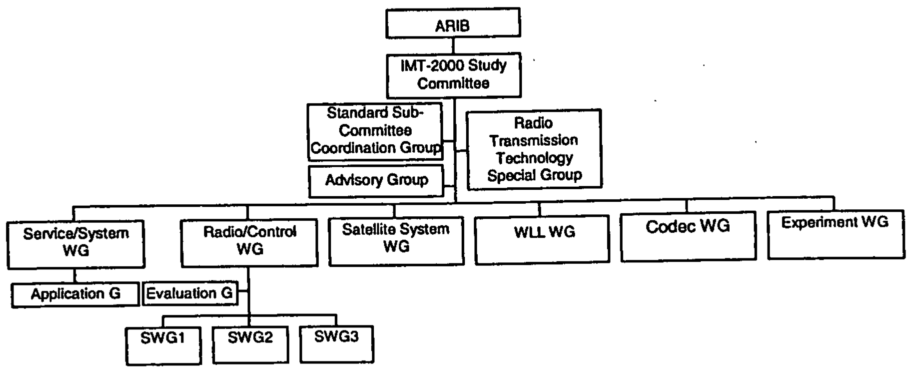


Figure 14.6 ARIB organization for IMT-2000.

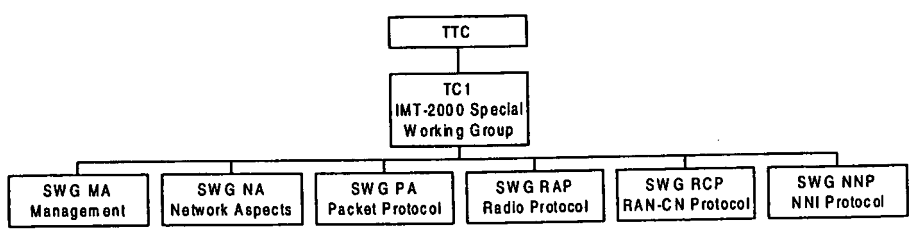


Figure 14.7 TTC organization for IMT-2000.

TTC, represented in Figure 14.7, is responsible for the development of system architecture, information flows, requirements for network control, layer 2 in cooperation with ARIB, layer 3, RAN-CN interface, network-to-network interface (NNI), call associated signaling, and interworking with other networks. Since Japan has chosen to base third generation network on GSM core network, close cooperation with ETSI is under way. For more information on the TTC standardization see [13].

14.2.5 Korea

In Korea, the government sets the policies for the mobile radio system standardization. For example, the adoption of IS-95 over other wireless technologies for Korean PCS was decided by the Korean government. The Korean goals for IMT-2000 are a global standard, global roaming, low cost terminal, to meet the ITU requirements, and backwards compatibility, if possible. Korea has a parallel approach to standards: international cooperation and development of domestic standards for implementation after 2002. The international cooperation includes relations with ARIB, ETSI, and TIA, as well as participation in the ITU RTT process.

The Telecommunication Technology Association (TTA) was established by the Korean Ministry of Communications (MIC) in 1988 to develop Korean Information Communication Standard (KICS) related to telecommunications. Within TTA, the IMT-2000 subcommittee has been established for the development of third generation standards. The structure of the subcommittee is depicted in Figure 14.8. For more information on the TTA standardization see [14].

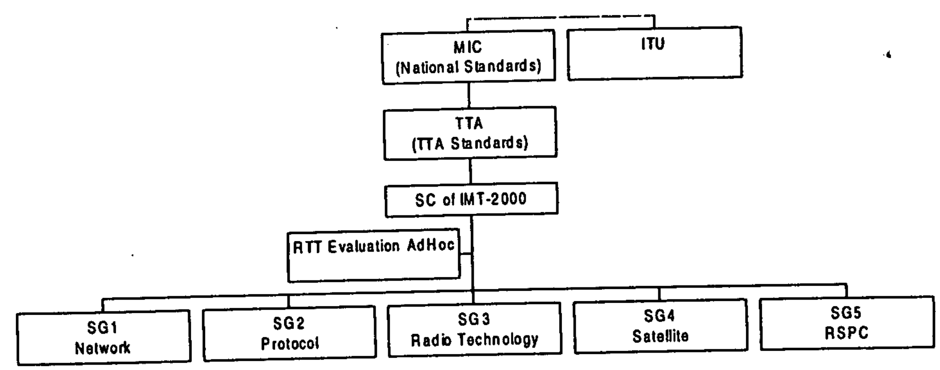


Figure 14.8 Structure of TTA.

The Electronics and Telecommunications Research Institute (ETRI) is an R&D organization developing technologies considered important for Korea. The Mobile Telecommunications Technology Division of ETRI is developing technology for IMT-2000. For more information on ETRI see [15].

14.2.6 China

In China, the Ministry of Post and Telecommunications (MPT) is driving IMT-2000 standardization. The Research Institute for Telecom & Transmission (RITT) is responsible for domestic standardization. It has actively cooperated with outside standardization bodies such as ETSI and ARIB. A MPT delegation has participated, for example, in the ETSI/SMG meetings as an associate member. MPT participates also actively the ITU process.

14.2.7 Other Groups

The Radio Standardization Meeting (RAST) is a group where different regional radio standardization organizations exchange views and plans regarding different areas of radio standardization, including third generation systems. The RAST activities are reported in [16].

The Future Advanced Mobile Wireless Universal System (FAMOUS) is a yearly meeting of administrations from Japan, the United States and the European Union. Its objectives are to provide a forum for discussion of matters related to the third generation mobile communications systems and to foster international cooperation regarding the interoperability of third generation systems on a global basis.

Asian-Pacific Wireless Forum (APWF) is a discussion forum for the purpose of Asian-Pacific cooperation in the R&D of wireless personal communications systems [17]. Based on the initiative of APWF, it has been proposed to set up a new standardization organization: Asian-Pacific Telecommunications Standards Institute (ATSI). Whether there is a need for an Asian-Pacific standard, and what the relation of the new organization in respect to the existing standards bodies would be, are still under discussion [17].

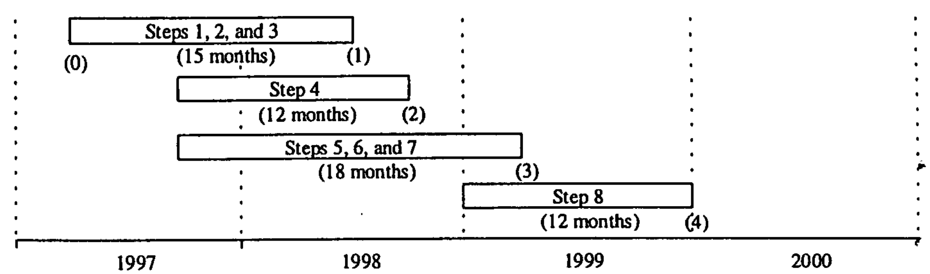
14.3 RADIO ACCESS NETWORK STANDARDIZATION

Both regional standards bodies (ETSI, ARIB, TIA, and TTA) and ITU-R TG8/1 have work programs to develop detailed specifications for the IMT-2000 radio access network. This raises a question: Which standards bodies are actually going to specify the IMT-2000 air interface and which are merely going to adopt already-made specifications? Since the ITU RTT evaluation process is aimed to produce a global standard, we will describe it and discuss its role with respect to regional standards activities.

Figures 14.9 and 14.10 illustrate the time schedule and the different steps of IMT-2000 air interface development, respectively. As can be seen, the consensus building within ITU starts only after evaluation documents have been submitted. However, the consensus building has already started based on initiatives from the regional standardization bodies and within the framework of FAMOUS meetings. In particular, Japan has established the CG group for harmonization purposes. In the beginning of 1998, TIA TR45.5 established an ad-hoc committee for IMT-2000

coordination and harmonization process of CDMA proposals. The role of the ITU process has thereby been to accelerate the regional developments towards closer cooperation.

Detailed standardization of the third generation systems, which are of an order of magnitude more complex than today's systems, will place a heavy burden on standards organizations. Not only do the new features need to be standardized, but also backwards compatibility aspects need to be taken into account. Although the parameters of some wideband CDMA proposals have been harmonized, maintenance of this harmonization during the standards process presents a major challenge. Small details need to be agreed on, and change requests will pour into standards bodies continuously. This high level harmonization can take place in ad-hoc groups, and to some extent, it has already happened. However, detailed standardization requires much more strict rules and framework. Having two or more standard bodies to do their detailed work separately and trying to achieve joint specification through ad-hoc meetings might prove to be very complex, if not impossible. Therefore, a joint forum for detailed standard development would be preferable.



Steps in the Radio Interface Development Process:

Step 1: ITU BR issues request for submission of candidate radio transmission technologies (RTTs)
Step 2: Manufacturers or other parties develop candidate RTTs
Step 3: Proponents submit candidate SRTT proposals
Step 4: Evaluation groups evaluate candidate RTTs
Step 5: TG 8/1 reviews evaluation activities
Step 6: TG 8/1 reviews candidate RTTs to ensure minimum performance capabilities are met
Step 7: TG 8/1 considers evaluation results, builds concensus, and determines the key technical characteristics of the IMT-2000 radio interface
Step 8: TG 8/1 develops radio interface specification Recommendations

Critical milestones in the Radio Interface Development Process:

(1) Issue request for candidate RTTs March 1997 ITU proposed cut off for submission of candidate RTT proposals, to be confirmed by evaluation groups June 1998
(2) Cut off for submission of evaluation reports to ITU September 1997
(3) TG 8/1 determines the key characteristics for the IMT-2000 radio interface March 1999
(4) TG 8/1 completes development of radio interface specification Recommendations December 1999 (approx.)

Figure 14.9 Time schedule of the radio interface development process (*Source*: [18], reproduced with permission of ITU).

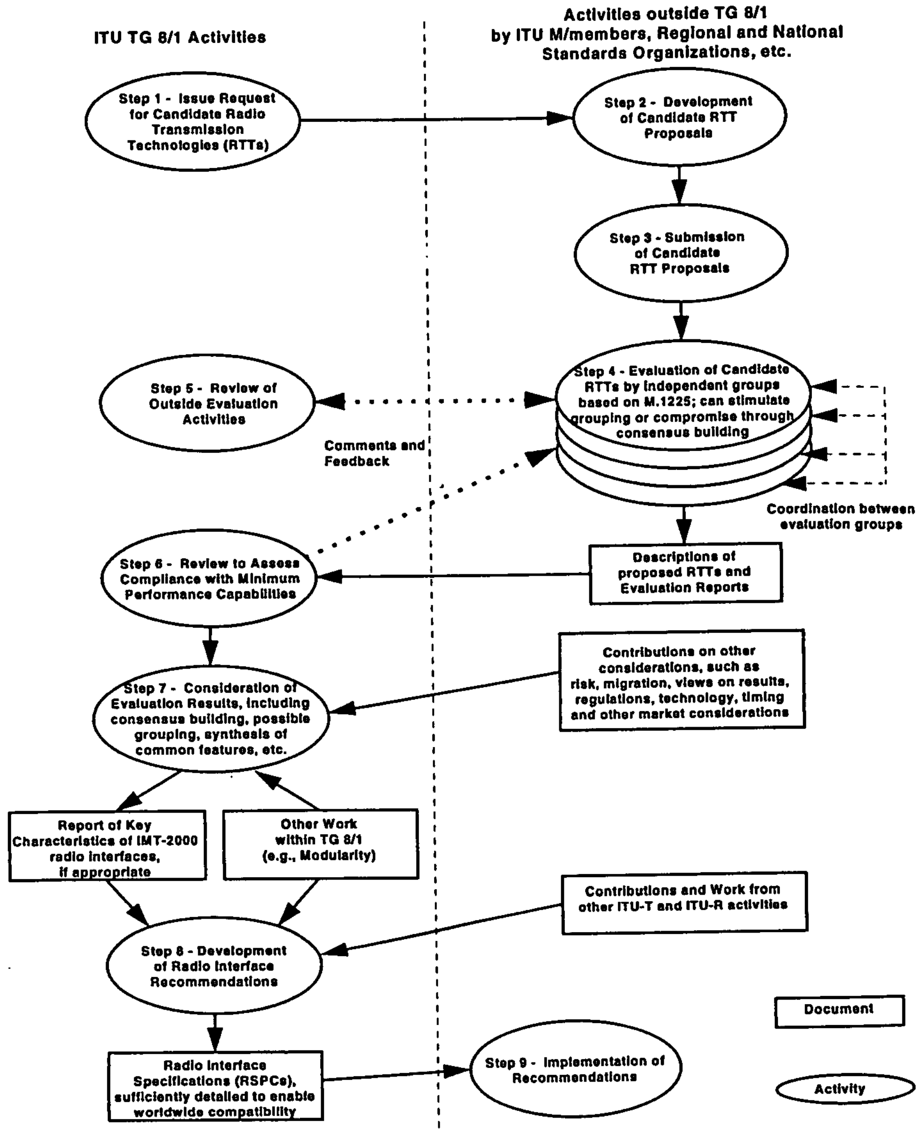


Figure 14.10 IMT-2000/FPLMTS radio interface development process (*Source*: [18], reproduced with permission of ITU)

One model for standards development is as follows. The core-standard would be developed by a joint standards body ensuring global conformance and minimum compatibility. Regional standard bodies would focus to specific issues related to their region's specific needs.

ITU activities encompass all third generation proposals. Given the current situation, where we have several proposals not all based on wideband CDMA, the ITU may not be flexible enough for this type of standards development. This seems to be the view of regional standards bodies, which are seeking cooperation outside ITU. Thus, a more likely development is that a global standard would be specified in a joint industry forum, for example.

14.3.1 ITU Evaluation Groups

The following evaluation groups have been formed [1]:

- IMT-2000 Evaluation Groups;
- Australia: Cooperative Research Centre for Broadband Telecommunications and Networking (CRC-BTN);
- Austria: ETSI Project DECT;
- Brazil: ANATEL;
- Canada: Canada Evaluation Group (CEG);
- China: China Evaluation Group (ChEG);
- Europe: ETSI SMG2, and INMARSAT;
- Japan: Association of Radio Industries & Businesses (ARIB);
- Korea: TTA Evaluation Ad-hoc;
- Malaysia: ITU;
- New Zealand: Radio Spectrum Management;
- USA:TIA/TR-45 Ad-hoc International Standards Development Group, and T1P1/TR46 International Standards Ad-Hoc;

14.3.2 IMT-2000 Radio Transmission Technology Proposals

The following proposals were submitted for terrestrial radio access by June 1998:

- ARIB/Japan: W-CDMA;
- CATT/China: TD-SCDMA;
- TTA Korea: Global CDMA I and II;
- ETSI project DECT: DECT;
- ETSI: UTRA;
- TIA/USA: TR45.3 (UWC-136) and TR45.5 (cdma2000), and TR46 (WIMS W-CDMA);
- T1P1-ATIS/USA: WCDMA/NA.

For satellite radio access the following proposals were submitted:

- European Space Agency (ESA): SW-CDMA and SW-CTDMA;
- ICO: ICO RTT;
- INMARSAT: Horizons;
- TTA Korea: satellite RTT.

14.4 CORE NETWORK STANDARDIZATION

The focus of third generation standards has been on the air interface. However, for a complete system the core network is equally important. There are two main views for the core network development. The first builds on a clean table approach where the core network would be based on B-ISDN. The second approach builds on the evolution from the second generation systems. Recently, the second approach has been gaining more support. For example, the adoption of the Family of Systems concept in ITU was a recognition of the evolution of second generation core networks.

The standardization of the evolving second generation core networks is driven by regional standardization bodies. ETSI SMG is responsible for GSM/UMTS core network standardization, and TIA TR45.2 for the IS-41 core network standardization. Japan has also decided to adopt the GSM/UMTS core network. Similar considerations as for the radio standardization apply for the core network standardization. A joint forum may need to be established for detailed specification development or one group must take a clear lead. The role of ITU-T is seen by many parties to be in global issues, such as common numbering and identifying interfaces required for interworking, rather than in developing detailed standards.

14.5 REGULATION

Fundamentally, regulatory matters are national issues. However, in some cases regional and global harmonization of regulations and policy issues is desirable [19]. Areas where common regulations may be required are licensing, services, provisioning, interconnection, infrastructure, frequencies, numbering, security, and policies [19]. With regard to a radio access system, the most important regulatory areas are spectrum coordination and licensing.

The process of obtaining frequencies can be divided into three phases: identification of spectrum, allocating the spectrum for specific purposes, and licensing the spectrum. Since the identified spectrum is usually already used for some other purposes, it needs to be cleaned from existing users before it can be used for the new purpose. Therefore, proactive long-term planning is required to ensure the availability of spectrum when required.

ITU plays an important role in spectrum regulation. The Radio Regulations (RR) of ITU are updated in World Radio Conference (WRC). The next WRC is scheduled for 1999. The national regulators are not bounded to follow the ITU guidelines for spectrum allocation. However, the ITU RRs form a tool to encourage them to do that in order to achieve global harmonization of spectrum [19]. The IMT-2000 spectrum was identified

in the year 1992 by WARC92 as a result of ITU studies on IMT-2000. These studies indicated that the minimum spectrum for IMT-2000 should be 230 MHz.

Spectrum is allocated for a specific purpose. It has to be decided what services can be provided, what technology is allowed to be used, and how the spectrum is allocated between operators. In addition to the IMT-2000 services, it might be desirable to provide some other services such as broadband data for indoor service within the IMT-2000 spectrum [20]. Whether this should be allowed or not must be decided. The spectrum use can be bounded to a specific standard, or the operator may be allowed to use any technology to deliver their services. The latter approach was taken in the US PCS spectrum assignments. The spectrum can be allocated on an exclusive basis or by using spectrum sharing. For an exclusive allocation, what is the minimum bandwidth for a single operator and what is the maximum number of operators that can be accommodated within the given frequency allocation?

The last phase in spectrum allocation is licensing. It has to be decided who can operate the networks. This can happen by "beauty contests" (i.e., who is evaluated to be most suitable to run a network), by lotteries, or by auctions. It has been argued whether the spectrum auction is a fair and effective way of assigning licenses for third generation mobile operators [20]. Furthermore, since spectrum is so scarce, it is a valuable asset. Thus, governments are eager to cash in that asset, as was seen in the United States for the PCS licenses. If it is decided to auction the spectrum, it has to be decided how auctions will be held and who can bid. The are three options for auctions: multiple round open auctions (used in the United States for the PCS frequencies), sealed bid auction, or royalty auction, where the payment takes form of a percentage of qualifying revenue. An important question is whether an existing operator can bid for the third generation license. This has to be weighted against the competitive advantage they might gain since they already have a nationwide coverage while the new operators have to build the coverage from scratch. The regulation could require national roaming between operators to alleviate the newcomers' position [20]. The regulation can also rule what the conditions are for a license. These include coverage requirements and the length of a license.

In general, Europe follows ITU recommendations for spectrum issues. The Europe-wide harmonization is carried out by the Conference of European Post and Telecommunications (CEPT). The European Commission can issue a directive to create harmonized frequency allocations for specific technologies. Examples of such directives are GSM in 1987 and DECT in 1991. Recently, the commission has published its proposed UMTS decision, which sets in place timescales and actions for national licences and spectrum harmonization by the year 2000 [21]. The European Radiocommunications Committee (ERC) of CEPT makes decisions which usually form the basis for the harmonized spectrum designations. CEPT has designated most of the IMT-2000 spectrum for UMTS with the adoption of the ERC Decision on UMTS, which identifies a total of 155 MHz of spectrum for terrestrial UMTS services, with an additional 60 MHz set aside for UMTS satellite services [22]. For the third generation frequencies, the United Kingdom intends to hold spectrum auctions in summer 1999. UMTS Auction Consultative Group (UACG) has been established to facilitate the UK government's ongoing consultation with industry and other interested parties on matters

relating to the proposed auction of spectrum licences for terrestrial UMTS (for more details, refer to website http://www.open.gov.uk/radiocom/rahome.htm).

In the United States, the Federal Communications Commission (FCC) is managing the spectrum issues. The Inter-American Telecommunication Commission (CITEL) is issuing recommendations on spectrum issues for North and South America. The recommendations of CITEL are nonbinding but carry significant weight. The third generation spectrum is for the most part already used by PCS systems. However, the United States is currently developing spectrum positions for the third generation including the identification of new frequency bands.

In Japan, the Ministry of Post and Telecommunication (MPT) is responsible for spectrum regulation. The Japanese spectrum plan follows the ITU recommendation for IMT-2000 frequencies.

14.6 MARKET CONSIDERATIONS

In this section, we discuss the wireless market development. Prediction of the future is always a difficult task, and the wireless communications field is not an exception to this rule. Long-term market projections for mobile radio systems are difficult when companies are struggling even with short-term projections. In this section we present some subscriber predictions for third generation systems in order to assess potential users at the time of the deployment of third generation systems.

There are considerable differences between the several market estimates for mobile users. The UMTS Forum has predicted that there would be 1.7 billion mobile users by 2010. This is 20% of the world's user population. The UMTS Forum also estimates that 45% of the mobile users will use high speed data services by 2010. It is clear, however, that there is no single right figure for the IMT-2000 market. The development of the third generation market depends on several factors including coverage, tariffing, lifestyle changes, and terminal offerings. Table 14.1 presents world subscriber growth for different compounded annual growth percentages for 1998-2001. According to [23], growth of 40 to 45 percent is considered most plausible. Between 1996 and 1997, the annual growth worldwide averaged 51.6 percent and since 1990, compounded annual growth has averaged 50.9 percent.

Table 14.1
Forecast of World Subscriber Growth (Millions)

	Assumed Compounded Annual Growth				
Year	25%	30%	35%	40%	45%
1998	258	268	279	289	299
1990	322	349	376	404	434
2000	403	453	508	566	629
2001	504	589	685	793	912

Source: [23].

The market for mobile multimedia services will depend on several factors. The market analysis group in the UMTS Forum has conducted a study that analyzed four different market scenarios for the mobile multimedia market [24]. These scenarios give an indication of the factors that will determine market development. A short summary of these scenarios follows.

Scenario 1, Slow Evolution: Mobile multimedia development is slow due to limited applications and high service and terminal prices. Unsuccessful liberalization has led to low competition. Fragmented standards have resulted in no economies of scale. In addition, fragmented standards hinder global application development. No global standard has been achieved, which has led to high equipment prices. Applications are difficult to reconfigure and to adapt to personal needs, and there have been no breakthrough developments for critical terminal technologies such as displays.

Scenario 2, Business Centric: Mobile multimedia takes off in the business sector but not in the consumer sector, as there is a lack of innovation in consumer applications. Lack of competition due to unsuccessful liberalization contributes to high premiums to access the service. However, high access costs can be afforded by business sector users. Terminals are intelligent and configurable to user needs but still difficult to use.

Scenario 3, Sophisticated Mass Market: Mass market for mobile multimedia has emerged. Terminals and applications have a large set of features and can be customized for personal needs. However, user interfaces are still rather complicated, being suitable for IT-literate users. Liberalization has resulted in high competition and thus low access premiums. Global standards for both systems and application platforms have facilitated low equipment costs and rich application developments. Spectrum was auctioned and the resulting costs are reflected in initially high prices that quickly drop.

Scenario 4, Commoditized Mass Market: A real mobile multimedia mass market has emerged and comprises both business and consumer users. Liberalization and adoption of global standards have resulted in economies of scale. Spectrum is cheap; simple and cheap terminals are available from the beginning.

Table 14.2 shows the number of mobile users and multimedia users by 2005 for these different scenarios. The essential differences between the first two and the last two scenarios concern the worldwide platform in both radio and traffic delivery and in application. The common platform facilitates economies of scale and sets a standardized basis for application development. The main difference between scenarios 1 and 2 is that in the second scenario users can configure their terminals according to their needs since they understand IT. In the first scenario, the user interface is more limited and thus not adequate even for most business users. The main difference between scenarios 3 and 4 is the higher initial cost in scenario 3 due to spectrum pricing and the price and ease of use for terminals. In scenario 3, spectrum is auctioned and terminals are bulky and expensive in the beginning. Furthermore, in scenario 3, tariffing is initially high due to

high investment into spectrum and equipment. These factors contribute to the slower development of mass market as with scenario 4.

Table 14.2
Number of Mobile and Multimedia Users in Europe by 2005

Scenario	Mobile users by 2005 (penetration)	Multimedia users by 2005
Slow evolution	82 M (22%)	7.5 M
Business centric	82 M (22%)	9 M
Sophisticated mass market	123 M (35%)	19 M
Commoditized mass market	140 M (40%)	27 M

Source:[24].

The potential new users for voice for data services can also be viewed from the perspective of the penetration rate by the year 2000. The higher the penetration rate of a service, the less potential there is for new business. On the other hand, a very high penetration rate might drive a deployment of a new system to support the required capacity.

Second generation systems are at the same time a threat and an opportunity for IMT-2000 systems. If data rates are enough for current applications and there is no lack of spectrum, the deployment of third generation systems will be slower. We should remember the predictions for analog systems in the beginning of the 1990s: they were supposed to be phased out by the middle of the decade. However, in 1998 AMPS has still a significant part of the cellular subscribers. We should always remember that we can learn from history, but we cannot predict the future based on history. In a chaotic process, even the smallest change in starting conditions will lead to a totally different end result.

Second generation systems are also tools to educate operators and customers for the data market. For example, operators can test the impact of tariffing on service demand, without the heavy investments required for third generation services. When users are already used to new types of services, adoption of data services will be easier. For many services, third generation will just mean better quality of service (i.e., faster service and higher reliability) when compared to services provided by second generation systems. Thus, second generation systems can be viewed as a logical stepping stone towards third generation systems, and certainly not something we need to discard as quickly as possible.

14.7 WIRELESS BROADBAND NETWORKS

Wireless broadband networks will provide user bit rates up to 20 Mbps, with a maximum range of some tens of meters. Wireless ATM is the most promising technology to implement wireless broadband networks. The basic idea of wireless ATM (WATM) is to extend the communications capabilities of wired ATM such as provision of different QoS classes to mobile users [25].

In Europe and the US, there is common allocation for mobile broadband systems in the 5-GHz frequency band. Similar allocation is also considered in Japan. The very

large amount of unlicensed spectrum facilitates simple air interface design because limited attention needs to be paid into the spectrum efficiency as in the cellular systems using licensed frequency bands. Furthermore, the tariffing structure for high speed data services could be different from the cellular since there is no price for the spectrum and allocating spectrum for data services does not decrease revenues from other services such as speech.

The 20-Mbps data rate and large allocations of unlicensed spectrum make the integration of WATM radio access with IMT-2000 an attractive opportunity to extend the current cellular business to a new market. The IMT-2000 would offer wide area mobility and coverage, while the WATM would offer hotspot broadband data services. This would require the integration of mobility management of IMT-2000 and WATM networks. This issue is considered at least in the ETSI Broadband Radio Access Network (BRAN) project [25].

14.8 SATELLITES

Satellites can provide coverage to remote, rural, and maritime areas where it would be too expensive or impossible to set up terrestrial infrastructure. Satellite-based systems providing mobile services are termed mobile satellite services (MSS) or satellite personal communications networks (S-PCN).

The planned S-PCN systems use nongeostationary orbits (Non-GEO). Non-GEOs can be subdivided into low earth orbit (LEO) and medium earth orbit (MEO), also called intermediate circular orbit (ICO), systems. LEO systems have altitudes from 800 to 1600 km and MEO systems around 10,000 km. Examples of non-GEO systems are Globalstar, Iridium, Loral, and Odyssey [26]. It should be noted that these systems are being deployed within the time frame of 1998–2002 (i.e., close to IMT-2000 deployment). The first services are speech and data up to 9.6 Kbps. The European third generation satellite activities are addressed in [27].

14.9 ENABLING TECHNOLOGIES

The applications and terminals need to be developed to a new level in order to utilize the extensive capabilities of third generation networks. The improvement in the processing power of DSPs and ASICs makes the radio part smaller and smaller, thus creating room for additional devices such as better displays, digital cameras, and many others. However, the increased software, higher number of application devices, and higher data call for improvements in battery technologies. Solar power technology might be used as a complementing energy source.

Digital cameras facilitate digitization of images. Although IMT-2000 will bring significant improvements to data rates, multimedia applications such as image transmission will require even higher bandwidths. Thus, efficient compression technologies are vital. Fractal coding is a new technology that could facilitate a breakthrough in this area [24].

Users viewing multimedia information will want to do it through a high-quality, high-resolution display. Mobile users use their terminals in very variable lighting conditions. Consequently, the display has to be capable of presenting the picture clearly regardless of external lighting. Of course, size is of key importance for portable devices.

In order to capitalize on mass market, user interfaces of wireless devices must be developed far beyond today's standards. Applications have to be easy to use, non-technical and understandable to a lay person. Voice recognition is one possible technique that can help with building user-friendly applications. Virtual reality is used to create a virtual environment for one user: a mobile user could imitate office conditions, for example, in a hotel room and could see the others in a realistic meting environment. Interactive virtual reality opens new possibilities for developing more attractive games that can be played against other users over the wireless link.

Context-aware applications are applications that change their behavior according to the user's present context—their location, who they are with, what is the time of the day, and so on [28]. In wireless systems, the context information can be used to either transfer information from the network to the user, or transfer information from the user to the network. One challenge to implement such applications is how to sense the user's context (location, surrounding people, time of the day, etc.). The Global Positioning System (GPS) is one way of locating users. However, it may not be accurate enough. Another possibility is to use the location information from the mobile radio system. The environment can be sensed by active badges, attached to people, identifying their carrier [28]. A standard way of creating context-aware applications (i.e., creating an actual program triggered by the context) has been presented in [28]. The methodology is based on a language similar to HTML used to create World Wide Web pages. The language is used to create *notes* that can be then executed based on the context. The context and notes can be passed from the user terminal to the backbone network and vice versa. An example would be when a user enters a certain area, notes for that area are passed from the network and then triggered by the terminal when arriving at specific locations within the given area. GSM's Short Message Service (SMS) has been used to pass context and notes between a GSM terminal and backbone network [28]. Some of the features for context-aware applications can be realized by the multicast feature of wireless networks (i.e., information is transmitted only within certain areas, as for price information in a supermarket). However, a more advanced environment for context-aware applications can take other factors than location into account, such as surrounding people, and thus, the applications can be tailored for individual users.

Operating systems form a platform for application development. EPOCH is an operating system for wireless information devices such as communicators and smart phones. Windows CE is an operating system developed for palmtop computers. The new version 2.0 adds wireless connectivity to this product. Sun Personal Java 1.0 is a similar type of product for designing application programming interfaces (API). The real value of these products are is that they create platforms for application development with which many are already familiar. These developments indicate that wireless Web services are going to be taken seriously. The key to mass market development is to create open interfaces for application developers.

A very significant development for creating wireless connectivity is the wireless application protocol (WAP), which aims for a global wireless network language and protocol stack. This would facilitate a common platform for development of applications for wireless Internet devices. The applications developed on top of WAP can run on future bearers and devices. The WAP Forum was created with the goal to make WAP an industry standard [29].

The new type of content places requirements for new types of charging mechanisms. Minute-based charging is not appropriate for data calls, but, for example, volume-based charging could be used. The importance of security will increase along with the number of users in wireless networks, as computer and wireless fraud are serious problems. Not only over-the-air security but also end-to-end security is important.

14.10 A WIRELESS DAY

Mike wakes up in the morning in Dallas and makes a video call to his wife Maureen, who is traveling and is on the other side of the globe in Tokyo. She answers from her hotel room and high-quality video connects them in a nice discussion. After the call, while getting ready for work, Mike wakes up his children, Amy and Matt, who want to first connect their Tamagotchi virtual pets with their friends to exchange vital information about their pets' dreams. This happens using improved real-time short message service, consuming hardly any bandwidth, and thus being cheap.

During his lunch break, Mike contacts his friend to play an interactive virtual reality video game through the wireless network. Bursts of images are transmitted through the wireless link. On the way back home Mike remembers that he needs to transmit his latest technical report to his boss. He stops on the road, connects to the corporate intranet, and a 10-Mbytes engineering document is transferred instantly through the IMT-2000 wireless connection. After driving a while, his car stops. What now, he wonders. There seems to be something wrong. He connects through video to his car service to get instructions. He focuses the digital camera in his wireless device into the car alarm lights. This view is transmitted to the car service, and based on it, they give him instructions on how to fix the problem.

Finally at home, Mike reads his email. Something from mom, he notices. A photograph taken in Hawaii, where she is on holiday. She took the photo with a digital camera and said, "send this to my son." Her IMT-2000 wireless terminal compressed the image, connected to the backbone network, and emailed the image to Mike over the wireless link.

14.11 CONCLUDING REMARKS

The industry has taken a big step forward to implement third generation systems. However, before commercial systems can emerge, an enormous effort is still required in standardization, regulation, and implementation. IMT-2000 has been the big business hype and will remain so for the next five years. Within that time span, we will see how

successful third generation will be in creating new business opportunities. However, human nature is such that part of the pioneers who were involved in setting up third generation standard frameworks will want to focus on new technologies and leave the details for other people. Thus, the next hype is already waiting for us beyond the horizon.

REFERENCES

[1] ITU website, http://www.itu.int/imt

[2] ETSI website, http://www.etsi.fr

[3] UMTS Task Force Report, February 21, 1996.

[4] The UMTS Forum website, http://www.umts-forum.org/

[5] GSM MoU website, http://www.gsmworld.com/

[6] Dimolitsas, S., "Standards Setting Bodies," *in The Mobile Communications Handbook*, J. D. Gibson (ed.), Boca Ratton: Florida: CRC Press, 1996.

[7] TIA website, http://www.tiaonline.org/

[8] T1 website, http://www.t1.org/

[9] CDG website, http://www.cdg.org

[10] UWCC website, htpp://www.uwcc.org

[11] Donegan, P., "Domestic Rivalries Unsettle Japan's Third Generation Push," *Mobile Communications International*, No. 40, April 1997, pp. 51–53.

[12] ARIB website, http://www.arib.or.jp/arib/english/index.html

[13] TTC website, http://www.ttc.or.jp/

[14] TTA website, http://www.tta.or.kr/

[15] ETRI website, http://www.etri.re.kr/

[16] RAST Website, http://www.rast.etsi.fi/

[17] Kohno, R., and M. Nagakawa, "International Cooperative Research and Development of Wireless Personal Communications in Asian-Pacific Countries", *IEEE Personal Communications*, Vol. 4, No. 2, April 1997, pp. 6–12.

[18] ITU Circular Letter 4 April 1997 (8/LCCE/47), "Request for Submission of Candidate Radio/Transmission Technologies (RTTs) for IMT 2000/FPLMTS Radio Interface".[1]

[19] Leite, F., R. Engelman, S. Kodama, H. Mennenga, and S. Towaij, "Regulatory Considerations Relating to IMT-2000," *IEEE Personal Communications*, August 1997, pp. 14–19.

[20] UK Department of Trade and Industry, "Mobile Communications on the Move," *A Consultation Document from the Department of Trade and Industry*, July 1997, can be found from website http://www.open.gov.uk/radiocom/rahome.htm.

[1] The complete volume of [18] can be obtained from ITU, General Secretariat – Sales and Marketing Service, Place des Nations - CH – 1211 GENEVA 20 (Switzerland), Tel. +41 22 730 61 41 (English)/ +41 22 730 61 42 (French), Fax. +41 22 730 51 94, Telex: 421 0000 uit ch, e-mail: sales@itu.int). It should be noted that the sole responsibility for selecting extracts from [18] for reproduction lies with the authors and can in no way attributed to ITU.

[21] European Commission, "Proposal by the Commission for a Decision of the European Parliament and Council on the co-ordinated introduction of mobile and wireless communications (UMTS) in the Community," 1997.

[22] "European Radiocommunications Committee ERC Decision of 30 June 1997 on the frequency bands for the introduction of the Universal Mobile Telecommunications System (UMTS)," ERC/DEC/(97)07 can be found from ERO website http://www.ero.dk/

[23] Shosteck, H., "The Explosion in World Cellular Growth", *The 2nd Annual UWC Global Summit*, Vancouver, BC, Canada, April 15-17, 1998.

[24] The UMTS Market Aspects Group, *UMTS Market Forecast Study*, 1997.

[25] Mikkonen, J., C. Corrado, C. Evci, and M. Prögler, "Emerging Wireless Broadband Networks," *IEEE Personal Communications*, Vol. 36, No. 2, February 1998, pp. 112–117.

[26] Comparetto, G., and R. Ramirez, "Trends in Mobile Satellite Technology," *IEEE Computer*, February 1997, pp. 44–52.

[27] Guntsch, A., M. Ibnkahla, G. Losquadro, M. Mazzella, D. Rovidas, and A. Timm, "EU's R&D Activities on Third-Generation Mobile Satellite Systems (S-UMTS)," *IEEE Personal Communications*, Vol. 36, No. 2, February 1998.

[28] Brown, P. J., J. D. Bovey, and X. Chen, "Context-Aware Applications: From the Laboratory to the Marketplace," *IEEE Personal Communications*, October 1997, pp. 58–63.

[29] WAP Forum website, http://www.wapforum.org

ABOUT THE AUTHORS

Tero Ojanperä was born in Korsnäs, Finland, on November 12, 1966. He received his M.Sc. degree from the University of Oulu, Finland, in 1991. He joined Nokia Mobile Phones, Oulu, Finland, in 1990 and worked as a research engineer from 1991 to 1992. From 1992 to 1995, he led a radio systems research group concentrating on wideband CDMA, GSM WLL, and US TDMA. From 1994 to 1995, he was also a project manager of the wideband CDMA concept development within Nokia. Later, this concept formed the basis for the FRAMES Wideband CDMA.

From 1995 to 1997, he was a research manager in Nokia Research Center, Helsinki, Finland, heading Nokia's third generation air interface research program, consisting of several air interface projects such as wideband CDMA, packet data, radio resource management algorithms, OFDM, and wideband TDMA. He was actively involved setting up the FRAMES project and during 1995-1996, he was also leader of the technical area air interface and the multiple access work package in the project, responsible for the selection of the FRAMES Multiple Access (FMA) scheme. The FRAMES Wideband CDMA was the basis for the UMTS WCDMA concept in ETSI. From 1994 to 1997, he was a Nokia representative for the UMTS radio interface issues in the ETSI SMG5 and SMG2 standardization committees.

From August 1997 to August 1998, he worked as a principal engineer in Nokia Research Center, Irving, TX. His was involved in the US third generation standards activities for the cdma2000. In addition, he was involved in technical/strategic work for Nokia's proposal for the UWC-136.

Since September 1998, he has been with Nokia Telecommunications, Finland, as a Head of Research, Radio Access Systems.

He has authored several conference and magazine papers as well as a chapter in two books *Wireless Communications TDMA vs. CDMA* (Kluwer Academic Publishers, 1997) and *GSM: Evolution towards 3rd generation* (Kluwer Academic Publishers, 1998). He has

three patents. During 1996-1997, he was a member of the Nokia patent committee and during 1997-1998, he was member of the patent committee of Nokia Mobile Phones. His research current interest lies in third generation wireless networks and multiuser detection.

Mr. Ojanperä is a member of the IEEE and during 1996 he was secretary of the IEEE Finland Section. He is a member of the International Advisory Committee of the Vehicular Technology Conference 1999 Fall.

Ramjee Prasad was born in Babhnaur (Gaya), Bihar, India on July 1, 1946. He is now a Dutch citizen. He received the B.Sc. (Eng.) from Bihar Institute of Technology, Sindri, India, and the M.Sc. (Eng.) and Ph.D. degrees from Birla Institute of Technology (BIT), Ranchi, India in 1968, 1970 and 1979, respectively.

He joined BIT as Senior Research Fellow in 1970 and became Associate Professor in 1980. While he was with BIT, he supervised many research projects in the area of Microwave and Plasma Engineering. During 1983-1988 he was with the University of Dar es Salaam (UDSM), Tanzania, where he became Professor in Telecommunications at the Department of Electrical Engineering in 1986. At UDSM, he was responsible for the collaborative project "Satellite Communications for Rural Zones" with Eindhoven University of Technology, The Netherlands. Since February 1988, he has been with the Telecommunications and Traffic Control Systems Group, Delft University of Technology (DUT), The Netherlands, where he is actively involved in the area of wireless personal and multimedia communications (WPMC). He is Head and Program Director of the Centre for Wireless and Personal Communications (CEWPC) of International Research Centre for Telecommunications-Transmission and Radar (IRCTR). He is currently involved in the European ACTS project FRAMES (Future Radio Wideband Multiple Access Systems) as a Project Leader of DUT. He has published over 200 technical papers as well as two books, *CDMA for Wireless Personal Communications* (Artech House, 1996) and *Universal Wireless Personal Communicatiions* (Artech House, 1998). His current research interest lies in wireless networks, packet communications, multiple access protocols, adaptive equalizers, spread-spectrum CDMA systems and multimedia communications.

He has served as a member of advisory and program committees of several IEEE international conferences. He has also presented keynote speeches, invited papers and tutorials on PIMRC at various universities, technical institutions and IEEE conferences. He was also a member of a working group of European cooperation in the field of scientific and technical research (COST-231) project dealing with "Evolution of Land Mobile Radio (including personal) Communications" as an expert for the Netherlands. He is now a member of the COST-259 project.

He is listed in the *US Who's Who in the World.* He was Organizer and Interim Chairman of IEEE Vehicular Technology/Communications Society Joint Chapter, Benelux Section. Presently, he is the elected chairman of the joint chapter. He is also the founder of the IEEE Symposium on Communications and Vehicular Technology (SCVT) in the Benelux and he was the Symposium Chairman of SCVT'93. He is the Co-ordinating Editor and Editor-in-Chief of a Kluwer international journal on "Wireless Personal Communications" and also a member of the editorial board of other international journals including *IEEE Communications Magazine* and *IEEE Electronics Communication Engineering Journal*. He was the Technical Program Chairman of PIMRC'94 International Symposium held in The Hague, The Netherlands during September 19-23, 1994 and also of the Third Communication Theory Mini-Conference in conjunction with GLOBECOM'94 held in San Francisco, California during November 27-30, 1994. He is the Conference Chairman of the 50th IEEE Vehicular Technology Conference to be held in Amsterdam, The Netherlands, during September 19-22, 1999. He is a Fellow of IEE, a Fellow of the Institution of Electronics & Telecommunication Engineers, a Senior Member IEEE and a Member of NERG (The Netherlands Electronics and Radio Society).

INDEX

Recent Titles in the Artech House Mobile Communications Series

John Walker, Series Editor

Advances in Mobile Information Systems, John Walker, editor

An Introduction to GSM, Siegmund M. Redl, Matthias K. Weber, Malcolm W. Oliphant

CDMA for Wireless Personal Communications, Ramjee Prasad

CDMA RF System Engineering, Samuel C. Yang

CDMA Systems Engineering Handbook, Jhong S. Lee and Leonard E. Miller

Cell Planning for Wireless Communications, Manuel F. Cátedra and Jesús Pérez-Arriaga

Cellular Communications: Worldwide Market Development, Garry A. Garrard

Cellular Mobile Systems Engineering, Saleh Faruque

The Complete Wireless Communications Professional: A Guide for Engineers and Managers, William Webb

GSM and Personal Communications Handbook, Siegmund M. Redl, Matthias K. Weber, Malcolm W. Oliphant

GSM Networks: Protocols, Terminology, and Implementation, Gunnar Heine

GSM System Engineering, Asha Mehrotra

Handbook of Land-Mobile Radio System Coverage, Garry C. Hess

Introduction to Radio Propagation for Fixed and Mobile Communications, John Doble

Introduction to Wireless Local Loop, William Webb

IS-136 TDMA Technology, Economics, and Services, Lawrence Harte, Adrian Smith, Charles A. Jacobs

Mobile Communications in the U.S. and Europe: Regulation, Technology, and Markets, Michael Paetsch

Mobile Data Communications Systems, Peter Wong, David Britland

Mobile Telecommunications: Standards, Regulation, and Applications, Rudi Bekkers and Jan Smits

Personal Wireless Communication With DECT and PWT, John Phillips, Gerard Mac Namee

RDS: The Radio Data System, Dietmar Kopitz, Bev Marks

RF and Microwave Circuit Design for Wireless Communications, Lawrence E. Larson, editor

Spread Spectrum CDMA Systems for Wireless Communications, Savo G. Glisic, Branka Vucetic

Understanding Cellular Radio, William Webb

Understanding Digital PCS: The TDMA Standard, Cameron Kelly Coursey

Understanding GPS: Principles and Applications, Elliott D. Kaplan, editor

Universal Wireless Personal Communications, Ramjee Prasad

Wideband CDMA for Third Generation Mobile Communications, Tero Ojanperä, Ramjee Prasad

Wireless Communications in Developing Countries: Cellular and Satellite Systems, Rachael E. Schwartz